Coleção
SCHAUM

Problema Resolvido

Álgebra Linear

Quarta edição

Mais de 600 problemas resolvidos

- Explicações concisas de todos os conceitos da disciplina
- Informação pertinente sobre sistemas algébricos, polinomiais e aplicações matriciais

COMPATÍVEL COM TODOS OS LIVROS-TEXTO!

Seymour Lipschutz e Marc Lipson

SEYMOUR LIPSCHUTZ é professor da Temple University e também já lecionou no Instituto Politécnico do Brooklyn. Recebeu seu Ph.D. em 1960 pelo Courant Institute da Universidade de Nova York. É um dos autores mais profícuos da Coleção Schaum. Em particular, entre outros, escreveu *Beginning Linear Algebra*, *Probability*, *Discrete Mathematics*, *Set Theory*, *Finite Mathematics* e *General Topology*.

MARC LARS LIPSON é professor da Universidade de Virgínia, tendo antes trabalhado na Universidade de Geórgia. É Ph.D. em Finanças desde 1994 pela Universidade de Michigan. Também é o coautor de *Matemática Discreta* e *Probability*, com Seymour Lipschutz.

L767a Lipschutz, Seymour.
Álgebra linear [recurso eletrônico] / Seymou Lipschutz, Marc Lars Lipson ; tradução: Dr. Claus Ivo Doering.–4.ed. – Dados eletrônicos. – Porto Alegre : Bookman, 2011.
(Coleção Schaum)

Editado também como livro impresso em 2011.
ISBN 978-85-407-0041-3

1. Matemática. 2. Álgebra linear. I. Lipson, Marc Lars. II. Título.

CDU 512

Catalogação na publicação: Ana Paula M. Magnus – CRB 10/2052

Seymour Lipschutz, Ph.D.

Marc Lars Lipson, Ph.D.

Álgebra Linear

Quarta edição

Tradução técnica

Dr. Claus Ivo Doering
Professor Titular do
Instituto de Matemática da UFRGS

Versão impressa
desta obra: 2011

2011

Prefácio

Nos últimos anos, a Álgebra Linear se tornou parte essencial do conhecimento matemático básico exigido de matemáticos e professores de Matemática, engenheiros, cientistas da computação, físicos, economistas e estatísticos, entre outros. Essa exigência reflete a importância e as múltiplas aplicações desse assunto.

Este livro foi desenvolvido para ser usado como livro-texto na disciplina de Álgebra Linear mas também pode ser usado como suplemento para outros livros. Apresenta uma introdução à Álgebra Linear que se mostrará útil a todos os leitores, independentemente de suas áreas de especialização. Incluiu-se mais material do que pode ser abordado na maioria dos cursos iniciais. Fizemos isso para tornar o livro mais flexível, para torná-lo um livro de referência útil e para estimular um maior desenvolvimento do material.

Cada capítulo começa com afirmações claras das definições, princípios e teoremas pertinentes, junto com material descritivo e ilustrativo adicional. A isso se segue um conjunto de exercícios graduais resolvidos e problemas complementares. Os problemas resolvidos servem para ilustrar e ampliar a teoria e fornecem a repetição de princípios básicos tão vital para o aprendizado. Várias demonstrações, especialmente as de todos os teoremas essenciais, estão incluídas entre os problemas resolvidos. Os problemas complementares servem como uma revisão completa do conteúdo de cada capítulo.

Nos três primeiros capítulos, tratamos de vetores no espaço euclidiano, álgebra de matrizes e sistemas de equações lineares. Esses capítulos proporcionam a motivação e as ferramentas computacionais básicas para as investigações abstratas de espaços vetoriais e transformações lineares que seguem. Depois de capítulos referentes a produto interno e ortogonalidade e determinantes, apresentamos uma discussão detalhada de autovalores e autovetores, dando condições para a representação de um operador linear por uma matriz diagonal. Isto leva naturalmente ao estudo das várias formas canônicas, especialmente a triangular, a de Jordan e a racional. Nos últimos capítulos, estudamos funcionais lineares e o espaço dual V^*, bem como formas bilineares, quadráticas e hermitianas. No último capítulo, tratamos de operadores lineares em espaços com produto interno.

As principais alterações desta quarta edição ocorreram nos apêndices. Expandimos o Apêndice A relativo a produtos tensorial e exterior de espaços vetoriais, incluindo demonstrações de existência e unicidade desses produtos. Acrescentamos apêndices relativos a estruturas algébricas, inclusive módulos e polinômios sobre um corpo e no Apêndice Miscelânea incluímos a inversa generalizada de Moore-Penrose, utilizada em várias aplicações, como na Estatística. Também introduzimos novos problemas resolvidos e complementares.

Finalmente, gostaríamos de agradecer à equipe da Coleção Schaum da McGraw-Hill, especialmente a Charles Wall, por sua constante cooperação.

SEYMOUR LIPSCHUTZ
MARC LARS LIPSON

Sumário

Vetores em $\mathbf{R}^n$ e $\mathbf{C}^n$, Vetores Espaciais

1.1 INTRODUÇÃO

A noção de vetor pode ser motivada ou por uma lista de números e índices, ou por meio de certos objetos da Física. Vejamos ambas maneiras.

Para isso, vamos supor que o leitor esteja familiarizado com as propriedades elementares do corpo dos números reais, denotado por **R**. Além desse corpo, vamos revisar algumas propriedades do corpo dos números complexos, denotado por **C**. No contexto de vetores, os elementos de nossos corpos numéricos são denominados *escalares*.

Mesmo que neste capítulo nos restrinjamos a vetores cujos elementos provenham de **R** e, mais tarde, de **C**, muitas das nossas operações também são aplicáveis a vetores cujas entradas sejam provenientes de algum corpo arbitrário K.

Lista de números

Digamos que os pesos (em kg) de oito universitários sejam dados pela lista

78, 63, 73, 62, 88, 73, 81, 97.

Utilizando apenas um símbolo, digamos, w, e índices subscritos distintos, podemos denotar os oito valores dessa lista, como segue.

$$w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8$$

Observe que cada índice denota a posição do valor na lista. Por exemplo,

$w_1 = 78$, o primeiro número, $w_2 = 63$, o segundo número da lista, ...

Uma lista de valores como essa,

$$w = (w_1, w_2, w_3, ..., w_8),$$

é denominada *tabela linear* ou *vetor*.

Vetores na Física

Muitas grandezas físicas, como temperatura e velocidade escalar, possuem apenas “magnitude”. Essas grandezas podem ser representadas por números reais e são denominadas *escalares*. Além dessas, também existem grandezas, como força e velocidade, que possuem tanto “magnitude” quanto “direção” e “sentido”. Essas grandezas são denominadas *vetores* e podem ser representadas por setas que começam em algum ponto referencial O dado, sendo dotadas de comprimento, direção e sentido apropriados.

Agora vamos supor que o leitor também esteja familiarizado com o espaço $\mathbf{R}^3$, em que todos os pontos são representados por ternos ordenados de números reais. Suponha que para o ponto referencial O mencionado esco-

(a) Soma de vetores   (b) Multiplicação por escalar

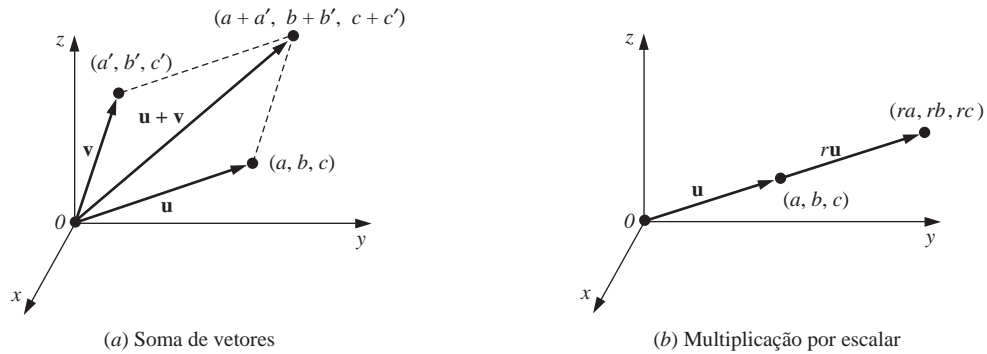



Figura 1-1

lhamos a origem dos eixos de $\mathbf{R}^3$. Então, cada vetor é determinado de maneira única pelas coordenadas de seu ponto final e vice-versa.

Existem duas operações importantes associadas aos vetores da Física, a soma de vetores e a multiplicação por escalar. Vejamos a definição dessas operações e as relações entre essas operações e os pontos finais dos vetores.

(i) ***Soma de vetores***. O vetor resultante $\mathbf{u} + \mathbf{v}$ de dois vetores $\mathbf{u}$ e $\mathbf{v}$ é obtido pela *lei do paralelogramo*, que diz que $\mathbf{u} + \mathbf{v}$ é a diagonal do paralelogramo formado por $\mathbf{u}$ e $\mathbf{v}$. Além disso, se (a, b, c) e (a', b', c') forem os pontos finais dos vetores $\mathbf{u}$ e $\mathbf{v}$, então o ponto final do vetor $\mathbf{u} + \mathbf{v}$ será dado por $(a + a', b + b', c + c')$. Essas propriedades podem ser vistas na Figura 1-1(a).

(ii) ***Multiplicação por escalar***. O múltiplo $r\mathbf{u}$ de um vetor $\mathbf{u}$ por um número real r é o vetor de mesma direção de $\mathbf{u}$ que é obtido multiplicando-se o tamanho de $\mathbf{u}$ por r e mantendo o mesmo sentido se $[r > 0]$ ou invertendo o sentido se $[r < 0]$. Além disso, se (a, b, c) for o ponto final do vetor $\mathbf{u}$, então (ra, rb, rc) é o ponto final do vetor $r\mathbf{u}$. Essas propriedades podem ser vistas na Figura 1-1(b).

Matematicamente, identificamos o vetor $\mathbf{u}$ com (a, b, c) e escrevemos $\mathbf{u} = (a, b, c)$. Além disso, o terno ordenado (a, b, c) de números reais é denominado ponto ou vetor, dependendo da interpretação. Generalizamos essa noção e dizemos que uma *ênupla* $(a_1, a_2, ..., a_n)$ de números reais é um vetor. Entretanto, para os vetores de $\mathbf{R}^3$, denominados vetores *espaciais*, podemos usar notação especial (Seção 1.6).

1.2 VETORES DE $\mathbf{R}^n$

O conjunto de todas as ênuplas de números reais, denotado por $\mathbf{R}^n$, é chamado de *espaço n-dimensional*. Uma ênupla específica de $\mathbf{R}^n$, digamos,

$$u = (a_1, a_2, ..., a_n),$$

é denominada *ponto* ou *vetor*. Os números a_i são denominados *coordenadas*, *componentes* ou *entradas* de u. Além disso, quando trabalhamos com o espaço $\mathbf{R}^n$, usamos o termo *escalar* para os elementos de $\mathbf{R}$.

Dizemos que dois vetores u e v são *iguais*, e escrevemos $u = v$, se possuírem o mesmo número de componentes e se os componentes correspondentes forem iguais. Embora os vetores (1, 2, 3) e (2, 3, 1) contenham os mesmos três números, esses vetores não são iguais, porque as entradas correspondentes não são iguais.

O vetor (0, 0, ..., 0), cujas entradas são todas 0, é denominado *vetor nulo*, ou *vetor zero*, e costuma ser denotado por 0.

Exemplo 1.1

(a) São vetores

$$(2, -5), \quad (7, 9), \quad (0, 0, 0) \quad \text{e} \quad (3, 4, 5).$$

Os dois primeiros pertencem a $\mathbf{R}^2$, enquanto os dois últimos pertencem a $\mathbf{R}^3$. O terceiro é o vetor nulo de $\mathbf{R}^3$.

(b) Encontre x, y e z tais que $(x - y, x + y, z - 1) = (4, 2, 3)$.

Pela definição de igualdade de vetores, os componentes correspondentes devem ser iguais. Assim,

$$x - y = 4, \qquad x + y = 2, \qquad z - 1 = 3$$

Resolvendo esse sistema de equações, obtemos $x = 4$, $y = -1$, $z = 4$.

Vetores coluna

Às vezes, escrevemos um vetor do espaço n-dimensional $\mathbf{R}^n$ verticalmente em vez de horizontalmente. Dizemos que um vetor desses é um *vetor coluna* e, nesse contexto, os vetores do Exemplo 1.1 escritos na forma horizontal são denominados *vetores linha*. Por exemplo, seguem vetores coluna com 2, 2, 3 e 3 componentes, respectivamente.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ -4 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 5 \\ -6 \end{bmatrix}, \quad \begin{bmatrix} 1{,}5 \\ \frac{2}{3} \\ -15 \end{bmatrix}$$

Também observamos que qualquer operação definida para vetores linha está definida para vetores coluna de maneira análoga.

1.3 SOMA DE VETORES E MULTIPLICAÇÃO POR ESCALAR

Considere dois vetores u e v de $\mathbf{R}^n$, digamos

$$u = (a_1, a_2, ..., a_n) \qquad \text{e} \qquad v = (b_1, b_2, ..., b_n)$$

A *soma* $u + v$ desses vetores é o vetor obtido somando os componentes correspondentes de u e v, ou seja,

$$u + v = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n)$$

A *multiplicação por escalar ru* do vetor u pelo número real r é o vetor obtido pela multiplicação de cada componente de u por r, ou seja,

$$ru = r(a_1, a_2, ..., a_n) = (ra_1, ra_2, ..., ra_n)$$

Observe que $u + v$ e ru também são vetores de $\mathbf{R}^n$. Não se define a soma entre vetores com número distinto de componentes.

O oposto de um vetor e a subtração de vetores em $\mathbf{R}^n$ são dados por

$$-u = (-1)u \qquad \text{e} \qquad u - v = u + (-v)$$

O vetor $-u$ é denominado *oposto* ou *negativo* de u e $u - v$ é a *diferença* de u e v.

Agora suponha que sejam dados os vetores $u_1, u_2, ..., u_m$ de $\mathbf{R}^n$ e os escalares $r_1, r_2, ..., r_m$ de $\mathbf{R}$. Podemos multiplicar os vetores pelos escalares correspondentes e então somar os múltiplos escalares resultantes para constituir o vetor

$$v = r_1u_1 + r_2u_2 + r_3u_3 + \cdots + r_mu_m$$

Um vetor v desses é denominado *combinação linear* dos vetores $u_1, u_2, ..., u_m$.

Exemplo 1.2

(a) Sejam $u = (2, 4, -5)$ e $v = (1, -6, 9)$. Então

$$\begin{aligned} u + v &= (2 + 1,\ 4 + (-5),\ -5 + 9) = (3, -1, 4) \\ 7u &= (7(2), 7(4), 7(-5)) = (14, 28, -35) \\ -v &= (-1)(1, -6, 9) = (-1, 6, -9) \\ 3u - 5v &= (6, 12, -15) + (-5, 30, -45) = (1, 42, -60) \end{aligned}$$

(b) O vetor zero $0 = (0, 0, ..., 0)$ de $\mathbf{R}^n$ é semelhante ao escalar 0 no seguinte sentido. Dado qualquer vetor $u = (a_1, a_2, ..., a_n)$, temos

$$u + 0 = (a_1 + 0, a_2 + 0, ..., a_n + 0) = (a_1, a_2, ..., a_n) = u$$

(c) Sejam $u = \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix}$ e $v = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$. Então $2u - 3v = \begin{bmatrix} 4 \\ 6 \\ -8 \end{bmatrix} + \begin{bmatrix} -9 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} -5 \\ 9 \\ -2 \end{bmatrix}$.

As propriedades básicas dos vetores sujeitos às operações de adição de vetores e de multiplicação por escalar são dadas no teorema a seguir.

Teorema 1.1 Dados quaisquer vetores u, v e w em $\mathbf{R}^n$ e escalares r e r' em $\mathbf{R}$,

(i)	$(u + v) + w = u + (v + w)$,	(v)	$r(u + v) = ru + rv$,
(ii)	$u + 0 = u$,	(vi)	$(r + r')u = ru + r'u$,
(iii)	$u + (-u) = 0$,	(vii)	$(rr')u = r(r'u)$,
(iv)	$u + v = v + u$,	(viii)	$1u = u$.

A demonstração do Teorema 1.1 será dada no Capítulo 2, onde aparece no contexto de matrizes (Problema 2.3).

Suponha que u e v sejam vetores de $\mathbf{R}^n$ tais que $u = rv$ para algum escalar não nulo r de $\mathbf{R}$. Então dizemos que u é um *múltiplo de* v. Dizemos, também, que u tem o *mesmo sentido* que v se $[r > 0]$ ou *sentido oposto* ao de v $[r < 0]$.

1.4 PRODUTO ESCALAR (OU INTERNO)

Considere dois vetores quaisquer u e v de $\mathbf{R}^n$, digamos,

$$u = (a_1, a_2, ..., a_n) \qquad \text{e} \qquad v = (b_1, b_2, ..., b_n)$$

O *produto escalar*, ou *produto interno*, de u e v é denotado e definido por

$$u \cdot v = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

Ou seja, $u \cdot v$ é obtido com a multiplicação dos componentes correspondentes e com a soma dos produtos resultantes. Dizemos que os vetores u e v são *ortogonais*, ou *perpendiculares*, se seu produto escalar for nulo, ou seja, se $u \cdot v = 0$.

Exemplo 1.3

(a) Sejam $u = (1, -2, 3)$, $v = (4, 5, -1)$ e $w = (2, 7, 4)$. Então

$$u \cdot v = 1(4) - 2(5) + 3(-1) = 4 - 10 - 3 = -9$$
$$u \cdot w = 2 - 14 + 12 = 0, \qquad v \cdot w = 8 + 35 - 4 = 39$$

Assim, u e w são ortogonais.

(b) Sejam $u = \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix}$ e $v = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$. Então $u \cdot v = 6 - 3 + 8 = 11$.

(c) Considere $u = (1, 2, 3, 4)$ e $v = (6, r, -8, 2)$. Encontre r tal que u e v sejam ortogonais.

Inicialmente obtemos $u \cdot v = 6 + 2r - 24 + 8 = -10 + 2r$. Faça então $u \cdot v = 0$ e resolva para r:

$$-10 + 2r = 0 \qquad \text{ou} \qquad 2r = 10 \qquad \text{ou} \qquad r = 5$$

Seguem as propriedades básicas do produto escalar de $\mathbf{R}^n$(mostradas no Problema 1.13).

Teorema 1.2 Dados quaisquer vetores u, v e w em $\mathbf{R}^n$ e escalar r em $\mathbf{R}$,

(i)	$(u + v) \cdot w = u \cdot w + v \cdot w$,	(iii)	$u \cdot v = v \cdot u$,
(ii)	$(ru) \cdot v = r(u \cdot v)$,	(iv)	$u \cdot u \geq 0$, se, e só se, $u = 0$.

Observe que (ii) afirma que "podemos tirar r para fora" da primeira posição de um produto escalar. A partir de (iii) e (ii), obtemos

$$u \cdot (rv) = (rv) \cdot u = r(v \cdot u) = r(u \cdot v)$$

Ou seja, também "podemos tirar r para fora" da segunda posição de um produto escalar.

O espaço $\mathbf{R}^n$ com essas operações de adição de vetores, multiplicação por escalar e produto escalar costuma ser denominado *espaço euclidiano n-dimensional*.

Norma (ou comprimento) de um vetor

A *norma* ou *comprimento* de um vetor u de $\mathbf{R}^n$, denotado por $\|u\|$, é definido como a raiz quadrada não negativa de $u \cdot u$. Em particular, se $u = (a_1, a_2, ..., a_n)$, então

$$\|u\| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

Ou seja, $\|u\|$ é a raiz quadrada da soma dos quadrados dos componentes de u. Assim, $\|u\| \geq 0$ e $\|u\| = 0$ se, e só se, $u = 0$.

Dizemos que um vetor u é *unitário* se $\|u\| = 1$ ou, de modo equivalente, se $u \cdot u = 1$. Dado qualquer vetor não nulo v de $\mathbf{R}^n$, o vetor

$$\hat{v} = \frac{1}{\|v\|} v = \frac{v}{\|v\|}$$

é o único vetor unitário de mesma direção e mesmo sentido de v. O processo de encontrar $\hat{v}$ a partir de v é denominado *normalização de* v.

Exemplo 1.4

(a) Seja $u = (1, -2, -4, 5, 3)$. Para obter $\|u\|$, podemos calcular primeiro $\|u\|^2$ tomando o quadrado de cada componente e somando, como segue.

$$\|u\|^2 = 1^2 + (-2)^2 + (-4)^2 + 5^2 + 3^2 = 1 + 4 + 16 + 25 + 9 = 55$$

Então $\|u\| = \sqrt{55}$.

(b) Sejam $v = (1, -3, 4, 2)$ e $w = (\frac{1}{2}, -\frac{1}{6}, \frac{5}{6}, \frac{1}{6})$. Então

$$\|v\| = \sqrt{1 + 9 + 16 + 4} = \sqrt{30} \qquad \text{e} \qquad \|w\| = \sqrt{\frac{9}{36} + \frac{1}{36} + \frac{25}{36} + \frac{1}{36}} = \sqrt{\frac{36}{36}} = \sqrt{1} = 1$$

Assim, w é um vetor unitário, mas não v. Contudo, podemos normalizar v.

$$\hat{v} = \frac{v}{\|v\|} = \left(\frac{1}{\sqrt{30}}, \frac{-3}{\sqrt{30}}, \frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}\right)$$

Esse é o único vetor unitário com a mesma direção e sentido de v.

A fórmula seguinte (que será demonstrada no Problema 1.14) é conhecida como a desigualdade de Schwarz, ou então, desigualdade de Cauchy-Schwarz. Ela é utilizada em muitos ramos da Matemática.

Teorema 1.3 (Schwarz) Dados quaisquer vetores u e v de $\mathbf{R}^n$, $|u \cdot v| \leq \|u\|\|v\|$.

Usando essa desigualdade, também demonstramos (Problema 1.15) o seguinte resultado, conhecido como "desigualdade triangular", ou desigualdade de Minkowski.

Teorema 1.4 (Minkowski) Dados quaisquer vetores u e v de $\mathbf{R}^n$, $\|u + v\| \leq \|u\| + \|v\|$.

Distância, ângulos e projeções

A *distância* entre os vetores $u = (a_1, a_2, ..., a_n)$ e $v = (b_1, b_2, ..., b_n)$ de $\mathbf{R}^n$ é denotada e definida por

$$d(u, v) = \|u - v\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

Pode ser mostrado que essa definição está de acordo com a noção usual de distância no plano euclidiano $\mathbf{R}^2$ ou no espaço $\mathbf{R}^3$.

O *ângulo* θ entre dois vetores não nulos u e v de $\mathbf{R}^n$ é definido por

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|}$$

O ângulo está bem definido porque, pela desigualdade de Schwartz (Teorema 1.3),

$$-1 \leq \frac{u \cdot v}{\|u\|\|v\|} \leq 1$$

Observe que se $u \cdot v = 0$, então $\theta = 90°$ (ou $\theta = \pi/2$). Isso está de acordo com nossa definição anterior de ortogonalidade.

A *projeção* de um vetor u sobre um vetor não nulo v é denotada e definida por

$$\text{proj}(u, v) = \frac{u \cdot v}{\|v\|^2} v = \frac{u \cdot v}{v \cdot v} v$$

Veremos adiante que essa definição está de acordo com a noção usual de projeção da Física.

Exemplo 1.5

(a) Suponha que $u = (1, -2, 3)$ e $v = (2, 4, 5)$. Então

$$d(u, v) = \sqrt{(1-2)^2 + (-2-4)^2 + (3-5)^2} = \sqrt{1 + 36 + 4} = \sqrt{41}$$

Para encontrar $\cos\theta$, onde θ é o ângulo entre u e v, primeiro determinamos

$$u \cdot v = 2 - 8 + 15 = 9, \qquad \|u\|^2 = 1 + 4 + 9 = 14, \qquad \|v\|^2 = 4 + 16 + 25 = 45$$

Então

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|} = \frac{9}{\sqrt{14}\sqrt{45}}$$

Também

$$\text{proj}(u, v) = \frac{u \cdot v}{\|v\|^2} v = \frac{9}{45}(2, 4, 5) = \frac{1}{5}(2, 4, 5) = \left(\frac{2}{5}, \frac{4}{5}, 1\right)$$

(b) Considere os vetores u e v dados na Figura 1-2(a) (com respectivos pontos finais em A e B). A projeção (perpendicular) de u sobre v é o vetor u^* de norma

$$\|u^*\| = \|u\| \cos\theta = \|u\| \frac{u \cdot v}{\|u\|\|v\|} = \frac{u \cdot v}{\|v\|}$$

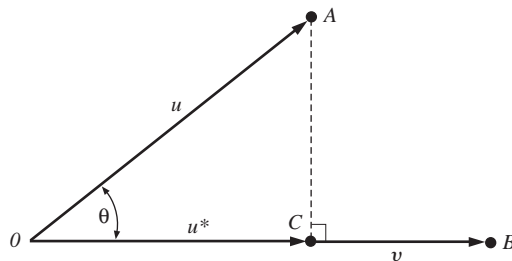
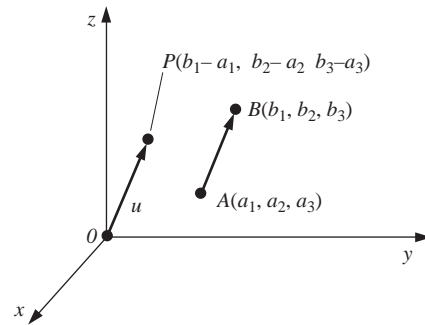



Figura 1-2

Para obter u^*, multiplicamos sua norma pelo vetor unitário na direção e sentido de v, obtendo

$$u^* = \|u^*\| \frac{v}{\|v\|} = \frac{u \cdot v}{\|v\|} \frac{v}{\|v\|} = \frac{u \cdot v}{\|v\|^2} v$$

Esse valor coincide com o dado na definição de proj(u, v).

1.5 VETORES APLICADOS, HIPERPLANOS, RETAS E CURVAS EM R^n

Nesta seção fazemos uma distinção entre a ênupla $P(a_i) \equiv P(a_1, a_2, \ldots, a_n)$ vista como um ponto de $\mathbf{R}^n$ e a ênupla $u = [c_1, c_2, ..., c_n]$ vista como uma seta (vetor) da origem O até o ponto $C(c_1, c_2, ..., c_n)$.

Vetores aplicados

Qualquer par de pontos $A(a_i)$ e $B(b_i)$ de $\mathbf{R}^n$ define um *vetor aplicado* ou um *segmento de reta orientado* de A para B, denotado por $\overrightarrow{AB}$. Identificamos o vetor $\overrightarrow{AB}$ com o vetor

$$u = B - A = [b_1 - a_1,\ b_2 - a_2,\ \ldots,\ b_n - a_n]$$

porque $\overrightarrow{AB}$ e u têm a mesma norma e a mesma direção. Isso está ilustrado na Figura 1-2(b) para os pontos $A(a_1, a_2, a_3)$ e $B(b_1, b_2, b_3)$ de $\mathbf{R}^3$ e o vetor $u = B - A$ de ponto final $P(b_1 - a_1, b_2 - a_2, b_3 - a_3)$.

Hiperplanos

Um *hiperplano* H em $\mathbf{R}^n$ é o conjunto de pontos $(x_1, x_2, ..., x_n)$ que satisfazem uma equação linear

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

em que o vetor de coeficientes $u = [a_1, a_2, ..., a_n]$ não é nulo. Assim, um hiperplano H em $\mathbf{R}^2$ é uma reta e um hiperplano H em $\mathbf{R}^3$ é um plano. Mostraremos adiante que, conforme mostrado na Figura 1-3(a) para o $\mathbf{R}^3$, u é ortogonal a qualquer segmento de reta orientado $\overrightarrow{PQ}$, em que $P(p_i)$ e $Q(q_i)$ são pontos de H. [Por essa razão, dizemos que u é *normal* a H e que H é normal a u.]

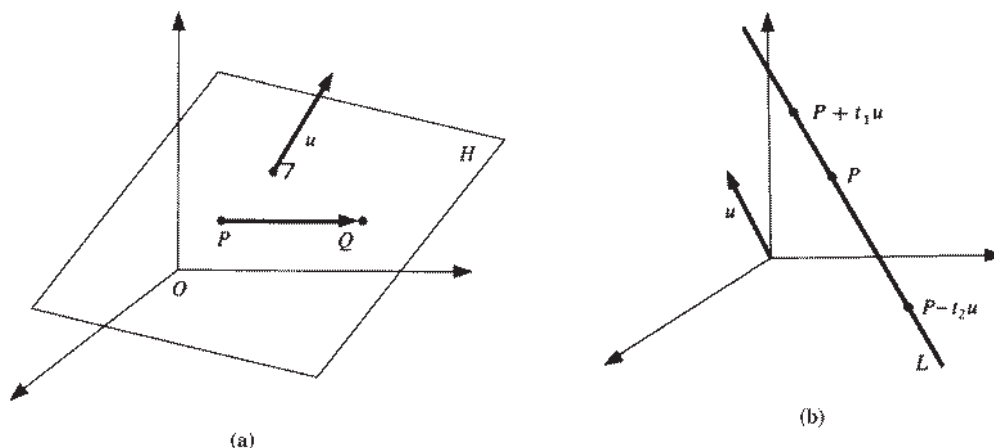



Figura 1-3

Como $P(p_i)$ e $Q(q_i)$ pertencem a H, esses pontos satisfazem a equação do hiperplano, ou seja,

$$a_1p_1 + a_2p_2 + \cdots + a_np_n = b \quad \text{e} \quad a_1q_1 + a_2q_2 + \cdots + a_nq_n = b$$

Seja $v = \overrightarrow{PQ} = Q - P = [q_1 - p_1, q_2 - p_2, \ldots, q_n - p_n]$

Então

$$u \cdot v = a_1(q_1 - p_1) + a_2(q_2 - p_2) + \cdots + a_n(q_n - p_n)$$
$$= (a_1q_1 + a_2q_2 + \cdots + a_nq_n) - (a_1p_1 + a_2p_2 + \cdots + a_np_n) = b - b = 0$$

Assim, $v = \overrightarrow{PQ}$ é ortogonal a u, como afirmado.

Retas em $\mathbf{R}^n$

A *reta* L de $\mathbf{R}^n$ que passa pelo ponto $P(b_1, b_2, ..., b_n)$ na direção e sentido do vetor não nulo $u = [a_1, a_2, ..., a_n]$ consiste nos pontos $X(x_1, x_2, ..., x_n)$ que satisfazem

$$X = P + tu \qquad \text{ou} \qquad \begin{cases} x_1 = a_1t + b_1 \\ x_2 = a_2t + b_2 \\ \dots\dots\dots\dots\dots \\ x_n = a_nt + b_n \end{cases} \qquad \text{ou} \quad L(t) = (a_it + b_i)$$

em que o *parâmetro t* percorre todos os valores reais. Uma tal reta L de $\mathbf{R}^3$ aparece na Figura 1-3(b).

Exemplo 1.6

(a) Seja H o plano de $\mathbf{R}^3$ que corresponde à equação linear $2x - 5y + 7z = 4$. Observe que $P(1, 1, 1)$ e $Q(5, 4, 2)$ são soluções dessa equação. Assim, P e Q, bem como o segmento de reta orientado

$$v = \overrightarrow{PQ} = Q - P = [5 - 1,\ 4 - 1,\ 2 - 1] = [4, 3, 1]$$

estão no plano H. O vetor $u = [2, -5, 7]$ é normal a H e, como era de se esperar,

$$u \cdot v = [2, -5, 7] \cdot [4, 3, 1] = 8 - 15 + 7 = 0$$

Portanto, u é ortogonal a v.

(b) Encontre uma equação do hiperplano H de $\mathbf{R}^4$ que passa pelo ponto $P(1, 3, -4, 2)$ e é normal ao vetor $v = [4, -2, 5, 6]$.

Os coeficientes das incógnitas na equação de H são os componentes do vetor normal u, portanto, a equação de H deve ser da forma

$$4x_1 - 2x_2 + 5x_3 + 6x_4 = k$$

Substituindo P nessa equação, obtemos

$$4(1) - 2(3) + 5(-4) + 6(2) = k \qquad \text{ou} \qquad 4 - 6 - 20 + 12 = k \qquad \text{ou} \qquad k = -10$$

Assim, a equação de H é dada por $4x_1 - 2x_2 + 5x_3 + 6x_4 = -10$.

(c) Encontre a representação paramétrica da reta L de $\mathbf{R}^4$ que passa pelo ponto $P(1, 2, 3, -4)$ e que tem a direção e o sentido de $u = [5, 6, -7, 8]$. Encontre, também, o ponto Q de L dado por $t = 1$.

Substituindo os componentes de P e u nos componentes da equação paramétrica geral dada, obtemos a representação paramétrica a seguir.

$$x_1 = 5t + 1, \qquad x_2 = 6t + 2, \qquad x_3 = -7t + 3, \qquad x_4 = 8t - 4$$

ou, equivalentemente,

$$L(t) = (5t + 1, 6t + 2, -7t + 3, 8t - 4)$$

Observe que usando $t = 0$ obtemos o ponto P de L. A substituição $t = 1$ fornece o ponto $Q(6, 8, -4, 4)$ de L.

Curvas em $\mathbf{R}^n$

Seja D um intervalo (finito ou não) da reta real $\mathbf{R}$. Uma função contínua $F: D \to \mathbf{R}^n$ é uma *curva* de $\mathbf{R}^n$. Assim, a cada ponto $t \in D$ está associado o ponto

$$F(t) = [F_1(t), F_2(t), ..., F_n(t)]$$

de [$\mathbf{R}^n$]. Além disso, a derivada de $F(t)$, se existir, fornece o vetor

$$V(t) = \frac{dF(t)}{dt} = \left[\frac{dF_1(t)}{dt}, \frac{dF_2(t)}{dt}, \ldots, \frac{dF_n(t)}{dt}\right]$$

que é tangente à curva. Normalizando $V(t)$, obtemos

$$\mathbf{T}(t) = \frac{V(t)}{\|V(t)\|}$$

Assim, $\mathbf{T}(t)$ é um vetor tangente unitário à curva. (Vetores unitários com significado geométrico costumam ser denotados em negrito.)

Exemplo 1.7 Considere a curva $F(t) = [\text{sen}\, t, \cos t, t]$ de $\mathbf{R}^3$. Tomando a derivada de $F(t)$ [ou, de cada componente de $F(t)$], obtemos

$$V(t) = [\cos t, -\text{sen}\, t, 1]$$

que é um vetor tangente à curva. Normalizemos $V(t)$. Inicialmente obtemos

$$\|V(t)\|^2 = \cos^2 t + \text{sen}^2\, t + 1 = 1 + 1 = 2$$

Decorre que o vetor tangente unitário $\mathbf{T}(t)$ dessa curva é dado por

$$\mathbf{T}(t) = \frac{V(t)}{\|V(t)\|} = \left[\frac{\cos t}{\sqrt{2}}, \frac{-\text{sen}\, t}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$$

1.6 VETORES DE R^3 (VETORES ESPACIAIS), NOTAÇÃO ijk

Os vetores de $\mathbf{R}^3$, denominados *vetores espaciais*, ocorrem em muitas aplicações, especialmente na Física. Existe até uma notação especial para três desses vetores, como segue.

$\mathbf{i} = [1, 0, 0]$ denota o vetor unitário na direção e sentido de x positivo.
$\mathbf{j} = [0, 1, 0]$ denota o vetor unitário na direção e sentido de y positivo.
$\mathbf{k} = [1, 0, 1]$ denota o vetor unitário na direção e sentido de z positivo.

Então cada vetor $u = [a, b, c]$ de $\mathbf{R}^3$ pode ser escrito de maneira única no formato

$$u = [a, b, c] = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

Como os vetores **i**, **j** e **k** são unitários e mutuamente ortogonais, obtemos os produtos escalares seguintes.

$$\mathbf{i}\cdot\mathbf{i} = 1, \quad \mathbf{j}\cdot\mathbf{j} = 1, \quad \mathbf{k}\cdot\mathbf{k} = 1 \qquad \text{e} \qquad \mathbf{i}\cdot\mathbf{j} = 0, \quad \mathbf{i}\cdot\mathbf{k} = 0, \quad \mathbf{j}\cdot\mathbf{k} = 0$$

Além disso, as operações vetoriais apresentadas anteriormente podem ser dadas em termos da notação **ijk** como segue. Suponha que

$$u = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \qquad \text{e} \qquad v = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

Então

$$u + v = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k} \qquad \text{e} \qquad ru = ra_1\mathbf{i} + ra_2\mathbf{j} + ra_3\mathbf{k}$$

onde r é um escalar. Também

$$u \cdot v = a_1b_1 + a_2b_2 + a_3b_3 \qquad \text{e} \qquad \|u\| = \sqrt{u \cdot u} = a_1^2 + a_2^2 + a_3^2$$

Exemplo 1.8 Suponha que $u = 3\mathbf{i} + 5\mathbf{j} - 2\mathbf{k}$ e $v = 4\mathbf{i} - 8\mathbf{j} + 7\mathbf{k}$.

(a) Para encontrar $u + v$, somamos componentes, obtendo $u + v = 7\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$

(b) Para encontrar $3u - 2v$, multiplicamos primeiro pelos escalares e depois somamos.

$$3u - 2v = (9\mathbf{i} + 13\mathbf{j} - 6\mathbf{k}) + (-8\mathbf{i} + 16\mathbf{j} - 14\mathbf{k}) = \mathbf{i} + 29\mathbf{j} - 20\mathbf{k}$$

(c) Para encontrar $u \cdot v$, multiplicamos os componentes correspondentes e depois somamos.

$$u \cdot v = 12 - 40 - 14 = -42$$

(d) Para encontrar $\|u\|$, tomamos a raiz quadrada da soma dos quadrados dos componentes.

$$\|u\| = \sqrt{9 + 25 + 4} = \sqrt{38}$$

Produto vetorial

Existe uma operação especial envolvendo vetores de $\mathbf{R}^3$ que não está definida em $\mathbf{R}^n$, com $n \neq 3$. Essa operação é denominada *produto vetorial* e é denotada por $u \times v$. Uma maneira de lembrar facilmente da fórmula para o cálculo de $u \times v$ é usar o determinante (de ordem dois) e seu simétrico, que são denotados e definidos como segue.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \qquad \text{e} \qquad -\begin{vmatrix} a & b \\ c & d \end{vmatrix} = bc - ad$$

Aqui, dizemos que a e d são os elementos da *diagonal* e b e c os elementos da *diagonal oposta*. Assim, o determinante é o produto ad dos elementos da diagonal menos o produto bc dos elementos da diagonal oposta, mas é o simétrico disso para o simétrico do determinante.

Supondo que $u = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ e $v = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, temos

$$\begin{aligned} u \times v &= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}\mathbf{i} \end{aligned}$$

Ou seja, os três componentes de $u \times v$ são obtidos da tabela

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

(que contém, na primeira linha, os componentes de u e, na segunda, os de v) como segue.

(1) Oculte a primeira coluna e calcule o determinante.
(2) Oculte a segunda coluna e calcule o simétrico do determinante.
(3) Oculte a terceira coluna e calcule o determinante.

Observe que $u \times v$ é um vetor, o que justifica a denominação de produto *vetorial*.

Exemplo 1.9 Encontre $u \times v$ nos casos seguintes. (a) $u = 4\mathbf{i} + 3\mathbf{j} + 6\mathbf{k}$, $v = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$; (b) $u = [2, -1, 5]$, $v = [3, 7, 6]$.

(a) Usamos $\begin{bmatrix} 4 & 3 & 6 \\ 2 & 5 & -3 \end{bmatrix}$ para obter $u \times v = (-9 - 30)\mathbf{i} + (12 + 12)\mathbf{j} + (20 - 6)\mathbf{k} = -39\mathbf{i} + 24\mathbf{j} + 14\mathbf{k}$.

(b) Usamos $\begin{bmatrix} 2 & -1 & 5 \\ 3 & 7 & 6 \end{bmatrix}$ para obter $u \times v = [-6 - 35, 15 - 12, 14 + 3] = [-41, 3, 17]$.

OBSERVAÇÃO O produto vetorial dos vetores **i**, **j** e **k** são os seguintes.

$$\begin{array}{lll} \mathbf{i} \times \mathbf{j} = \mathbf{k}, & \mathbf{j} \times \mathbf{k} = \mathbf{i}, & \mathbf{k} \times \mathbf{i} = \mathbf{j} \\ \mathbf{j} \times \mathbf{i} = -\mathbf{k}, & \mathbf{k} \times \mathbf{j} = -\mathbf{i}, & \mathbf{i} \times \mathbf{k} = -\mathbf{j} \end{array}$$

Assim, se interpretarmos o terno (**i**, **j**, **k**) como uma permutação cíclica, em que **i** segue **k** e, portanto, **k** antecede **i**, então o produto vetorial de dois deles no sentido dado é o terceiro, mas o produto vetorial de dois deles no sentido oposto é o simétrico do terceiro.

Duas propriedades importantes do produto vetorial estão destacadas no próximo teorema.

Volume $= u \cdot v \times w$
(a)

Plano complexo
(b)

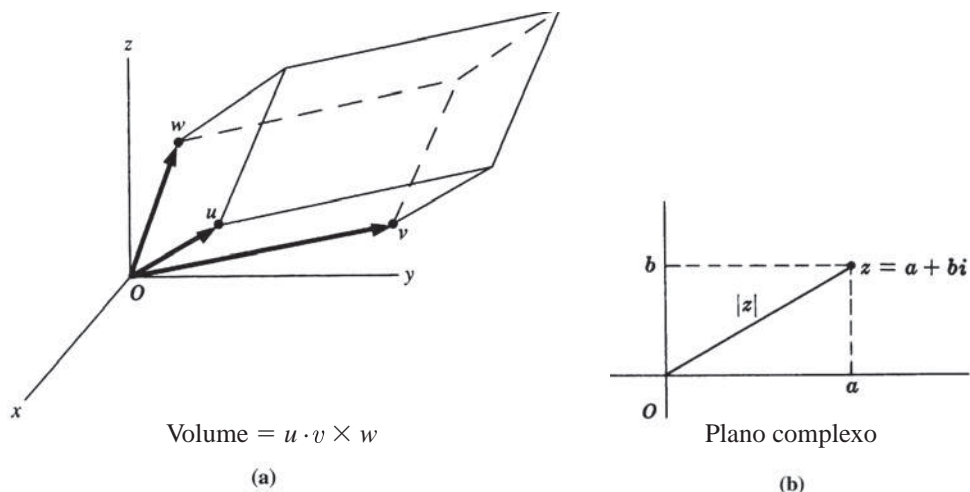



Figura 1-4

Teorema 1.5 Sejam u, v e w vetores de $\mathbf{R}^3$.

(a) O vetor $u \times v$ é ortogonal a ambos, u e v.

(b) O valor absoluto do "produto triplo"

$$u \cdot v \times w$$

representa o volume do paralelepípedo formado pelos vetores u, v e w. [Ver Figura 1-4(a).]

Observamos que os vetores u, v e $u \times v$ formam um sistema orientado de mão direita e vale a fórmula seguinte para o módulo de $u \times v$,

$$\|u \times v\| = \|u\|\|v\| \operatorname{sen}\theta$$

onde θ é o ângulo entre u e v.

1.7 NÚMEROS COMPLEXOS

O conjunto dos números complexos é denotado por **C**. Formalmente, um número complexo é um par ordenado (a, b) de números reais, sendo a igualdade, a adição e a multiplicação desses pares definidas como segue.

$$(a, b) = (c, d) \quad \text{se, e só se,} \quad a = c \quad \text{e} \quad b = d$$
$$(a, b) + (c, d) = (a + c, \ b + d)$$
$$(a, b) \cdot (c, d) = (ac - bd, \ ad + bc)$$

Identificamos o número real a com o complexo $(a, 0)$, ou seja,

$$a \leftrightarrow (a, 0)$$

Isso é possível porque as operações de adição e multiplicação de números reais são preservadas por essa correspondência, isto é,

$$(a, 0) + (b, 0) = (a + b, \ 0) \qquad \text{e} \qquad (a, 0) \cdot (b, 0) = (ab, 0)$$

Assim, vemos **R** como um subconjunto de **C** e substituímos $(a, 0)$ por a sempre que for conveniente e possível.

Observamos que o conjunto **C** dos números complexos, dotado das operações de adição e multiplicação assim definidas, constitui um *corpo* numérico, da mesma forma que o conjunto **R** dos números reais e o conjunto **Q** dos números racionais.

O número complexo (0, 1) é denotado por i. Esse número tem a importante propriedade de que

$$i^2 = ii = (0,1)(0,1) = (-1,0) = -1 \qquad \text{ou} \qquad i = \sqrt{-1}$$

Dessa forma, qualquer número complexo $z = (a, b)$ pode ser escrito da forma

$$z = (a,b) = (a,0) + (0,b) = (a,0) + (b,0)\cdot(0,1) = a + bi$$

Essa notação $z = a + bi$, em que $a \equiv \operatorname{Re} z$ e $b \equiv \operatorname{Im} z$ são denominadas, respectivamente, a *parte real* e a *parte imaginária* de z, é mais conveniente do que a de par ordenado (a, b). Isso se deve ao fato de que a soma e o produto dos números complexos $z = a + bi$ e $w = c + di$ podem ser efetuados simplesmente usando a comutatividade e a distributividade e lembrando que $i^2 = -1$, como segue.

$$\begin{aligned} z + w &= (a+bi) + (c+di) = a + c + bi + di = (a+b) + (c+d)i \\ zw &= (a+bi)(c+di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i \end{aligned}$$

Também definimos o *simétrico* de z e a subtração em **C** por

$$-z = -1z \qquad \text{e} \qquad w - z = w + (-z)$$

Atenção: A letra i que representa $\sqrt{-1}$ não tem relação alguma com o vetor $\mathbf{i} = [1, 0, 0]$ da Seção 1.6.

Conjugado complexo, valor absoluto

Considere o número complexo $z = a + bi$. O *conjugado* de z é denotado e definido por

$$\bar{z} = \overline{a + bi} = a - bi$$

Então $z\bar{z} = (a + bi)(a - bi) = a^2 - b^2 i^2 = a^2 + b^2$. Observe que z é real se, e só se, $\bar{z} = z$.

O *valor absoluto* de z, denotado por $|z|$, é definido como a raiz quadrada não negativa de $z\bar{z}$, ou seja, por

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

Observe que $|z|$ é igual à norma do vetor (a, b) de $\mathbf{R}^2$.

Suponha que $z \neq 0$. Então o inverso z^{-1} de z e a divisão em **C** de w por z são definidos, respectivamente, por

$$z^{-1} = \frac{\bar{z}}{z\bar{z}} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i \qquad \text{e} \qquad \frac{w}{z} = \frac{w\bar{z}}{z\bar{z}} = wz^{-1}$$

Exemplo 1.10 Suponha que $z = 2 + 3i$ e $w = 5 - 2i$. Então

$$\begin{aligned} z + w &= (2 + 3i) + (5 - 2i) = 2 + 5 + 3i - 2i = 7 + i \\ zw &= (2 + 3i)(5 - 2i) = 10 + 15i - 4i - 6i^2 = 16 + 11i \\ \bar{z} &= \overline{2 + 3i} = 2 - 3i \qquad \text{e} \qquad \bar{w} = \overline{5 - 2i} = 5 + 2i \\ \frac{w}{z} &= \frac{5 - 2i}{2 + 3i} = \frac{(5 - 2i)(2 - 3i)}{(2 + 3i)(2 - 3i)} = \frac{4 - 19i}{13} = \frac{4}{13} - \frac{19}{13}i \\ |z| &= \sqrt{4 + 9} = \sqrt{13} \qquad \text{e} \qquad |w| = \sqrt{25 + 4} = \sqrt{29} \end{aligned}$$

Plano complexo

O conjunto **R** dos números reais pode ser representado pelos pontos de uma reta. Analogamente, o conjunto **C** dos números complexos pode ser representado pelos pontos de um plano. Mais precisamente, consideramos que o ponto (a, b) do plano representa o número complexo $z = a + bi$, conforme Figura 1-4 (b). Nesse caso, $|z|$ é a distância da origem O do plano ao ponto z. O plano com essa representação é denominado *plano complexo*, assim como a reta que representa **R** é denominada *reta real*.

1.8 VETORES DE C^n

O conjunto de todas as ênuplas de números complexos, denotado por $\mathbf{C}^n$, é denominado *espaço n-dimensional complexo*. Assim como no caso real, os elementos de $\mathbf{C}^n$ são denominados *pontos* ou *vetores*, os elementos de **C** são os *escalares* e a soma de vetores em $\mathbf{C}^n$ e a multiplicação por escalar em $\mathbf{C}^n$ são dadas por

$$[z_1, z_2, \ldots, z_n] + [w_1, w_2, \ldots, w_n] = [z_1 + w_1,\ z_2 + w_2,\ \ldots,\ z_n + w_n]$$
$$z[z_1, z_2, \ldots, z_n] = [zz_1, zz_2, \ldots, zz_n]$$

com z_k, w_k e z em **C**.

Exemplo 1.11 Considere os vetores $u = [2 + 3i, 4 - i, 3]$ e $v = [3 - 2i, 5i, 4 - 6i]$ de $\mathbf{C}^3$. Então

$$u + v = [2 + 3i,\ 4 - i,\ 3] + [3 - 2i,\ 5i,\ 4 - 6i] = [5 + i,\ 4 + 4i,\ 7 - 6i]$$
$$(5 - 2i)u = [(5 - 2i)(2 + 3i),\ (5 - 2i)(4 - i),\ (5 - 2i)(3)] = [16 + 11i,\ 18 - 13i, 15 - 6i]$$

Produto escalar (ou interno) em C^n

Considere os vetores $u = [z_1, z_2, ..., z_n]$ e $v = [w_1, w_2, ..., w_n]$ em $\mathbf{C}^n$. O *produto escalar*, ou *produto interno*, de u e v é denotado e definido por

$$u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

Essa definição reduz ao caso real, pois $\bar{w}_K = w_K$ quando w_k for real. A norma de u é definida por

$$\|u\| = \sqrt{u \cdot u} = \sqrt{z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

Reforçamos que $u \cdot u$ e, portanto, $\|u\|$ são reais e positivos quando $u \neq 0$ e iguais a 0 quando $u = 0$.

Exemplo 1.12 Considere os vetores $u = [2 + 3i, 4 - i, 3 + 5i]$ e $v = [3 - 4i, 5i, 4 - 2i]$ de $\mathbf{C}^3$. Então

$$u \cdot v = (2 + 3i)(\overline{3 - 4i}) + (4 - i)(\overline{5i}) + (3 + 5i)(\overline{4 - 2i})$$
$$= (2 + 3i)(3 + 4i) + (4 - i)(-5i) + (3 + 5i)(4 + 2i)$$
$$= (-6 + 13i) + (-5 - 20i) + (2 + 26i) = -9 + 19i$$
$$u \cdot u = |2 + 3i|^2 + |4 - i|^2 + |3 + 5i|^2 = 4 + 9 + 16 + 1 + 9 + 25 = 64$$
$$\|u\| = \sqrt{64} = 8$$

O espaço $\mathbf{C}^n$, com as operações dadas de soma de vetores, multiplicação por escalar e produto interno, é denominado *espaço euclidiano n-dimensional complexo*. O Teorema 1.2 dado para $\mathbf{R}^n$ também é válido em $\mathbf{C}^n$, bastando substituir $u \cdot v = v \cdot u$ por

$$u \cdot v = \overline{u \cdot v}$$

Além disso, as desigualdades de Schwarz (Teorema 1.3) e de Minkowski (Teorema 1.4) são verdadeiras em $\mathbf{C}^n$ sem alteração alguma.

Problemas Resolvidos

Vetores de R^n

1.1 Decida quais vetores dentre os dados são iguais.

$u_1 = (1, 2, 3)$, $u_2 = (2, 3, 1)$, $u_3 = (1, 3, 2)$, $u_4 = (2, 3, 1)$

Dois vetores são iguais se suas entradas correspondentes forem iguais; logo, só $u_2 = u_4$.

1.2 Sejam $u = (2, -7, 1)$, $v = (-3, 0, 4)$ e $w = (0, 5, -8)$. Encontre

(a) $3u - 4v$,

(b) $2u + 3v - 5w$.

Primeiro multiplicamos pelos escalares e só depois efetuamos a soma dos vetores.

(a) $3u - 4v = 3(2, -7, 1) - 4(-3, 0, 4) = (6, -21, 3) + (12, 0, -16) = (18, -21, -13)$

(b) $2u + 3v - 5w = (4, -14, 2) + (-9, 0, 12) + (0, -25, 40) = (-5, -39, 54)$

1.3 Sejam $u = \begin{bmatrix} 5 \\ 3 \\ -4 \end{bmatrix}$, $v = \begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix}$ e $w = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$. Encontre

(a) $5u - 2v$,

(b) $-2u + 4v - 3w$.

Primeiro multiplicamos pelos escalares e só depois efetuamos a soma dos vetores.

(a) $5u - 2v = 5\begin{bmatrix} 5 \\ 3 \\ -4 \end{bmatrix} - 2\begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 25 \\ 15 \\ -20 \end{bmatrix} + \begin{bmatrix} 2 \\ -10 \\ -4 \end{bmatrix} = \begin{bmatrix} 27 \\ 5 \\ -24 \end{bmatrix}$

(b) $-2u + 4v - 3w = \begin{bmatrix} -10 \\ -6 \\ 8 \end{bmatrix} + \begin{bmatrix} -4 \\ 20 \\ 8 \end{bmatrix} + \begin{bmatrix} -9 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} -23 \\ 17 \\ 22 \end{bmatrix}$

1.4 Encontre x e y tais que valham as igualdades: (a) $(x, 3) = (2, x + y)$, (b) $(4, y) = x(2, 3)$.

(a) Como os vetores são iguais, igualamos seus componentes correspondentes, obtendo

$$x = 2, \qquad 3 = x + y$$

Resolvendo as equações lineares, obtemos $x = 2$, $y = 1$.

(b) Multiplicando pelo escalar x obtemos $(4, y) = (2x, 3x)$. Igualando as entradas correspondentes, obtemos

$$4 = 2x, \qquad y = 3x$$

Resolvendo as equações, obtemos $x = 2$, $y = 6$.

1.5 Escreva o vetor $v = (1, -2, 5)$ como uma combinação linear dos vetores $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 3)$ e $u_3 = (2, -1, 1)$.

Queremos escrever v no formato $v = xu_1 + yu_2 + zu_3$, com x, y e z ainda desconhecidos. Inicialmente, temos

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = x\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + z\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ x + 2y - z \\ x + 3y + z \end{bmatrix}$$

(É mais conveniente escrever vetores como colunas, e não como linhas, quando formamos combinações lineares.) Igualando as entradas correspondentes, obtemos

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

A única solução desse sistema triangular é $x = -6$, $y = 3$ e $z = 2$. Assim, $v = -6u_1 + 3u_2 + 2u_3$.

1.6 Escreva $v = (2, -5, 3)$ como uma combinação linear de

$$u_1 = (1, -3, 2),\ u_2 = (2, -4, -1) \text{ e } u_3 = (1, -5, 7).$$

Obtenha o sistema de equações lineares equivalente e resolva. Temos

$$\begin{bmatrix} 2 \\ -5 \\ 3 \end{bmatrix} = x\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ -4 \\ -1 \end{bmatrix} + z\begin{bmatrix} 1 \\ -5 \\ 7 \end{bmatrix} = \begin{bmatrix} x + 2y + z \\ -3x - 4y - 5z \\ 2x - y + 7z \end{bmatrix}$$

Igualando as entradas correspondentes, obtemos

$$\begin{aligned} x + 2y + z &= 2 \\ -3x - 4y - 5z &= -5 \\ 2x - y + 7z &= 3 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ -5y + 5z &= -1 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ 0 &= 3 \end{aligned}$$

A terceira equação $0x + 0y + 0z = 3$ indica que esse sistema não tem solução. Assim, v não pode ser escrito como uma combinação linear dos vetores u_1, u_2 e u_3.

Produto escalar (interno), ortogonalidade, norma em $\mathbf{R}^n$

1.7 Obtenha $u \cdot v$, sendo

(a) $u = (2, -5, 6)$ e $v = (8, 2, -3)$,

(b) $u = (4, 2, -3, 5, -1)$ e $v = (2, 6, -1, -4, 8)$.

Multiplicamos os componentes correspondentes e somamos.

(a) $u \cdot v = 2(8) - 5(2) + 6(-3) = 16 - 10 - 18 = -12$

(b) $u \cdot v = 8 + 12 + 3 - 20 - 8 = -5$

1.8 Sejam $u = (5, 4, 1)$, $v = (3, -4, 1)$ e $w = (1, -2, 3)$. Quais pares desses vetores (se houver) são perpendiculares (ortogonais)?

Calculamos o produto escalar de cada par de vetores.

$$u \cdot v = 15 - 16 + 1 = 0, \qquad v \cdot w = 3 + 8 + 3 = 14, \qquad u \cdot w = 5 - 8 + 3 = 0$$

Assim, u e v são ortogonais, u e w são ortogonais, mas v e w não são.

1.9 Obtenha k tal que u e v sejam ortogonais, nos casos seguintes.

(a) $u = (1, k, -3)$ e $v = (2, -5, 4)$,

(b) $u = (2, 3k, -4, 1, 5)$ e $v = (6, -1, 3, 7, 2k)$.

Calculamos $u \cdot v$, igualamos a zero e resolvemos em k.

(a) $u \cdot v = 1(2) + k(-5) - 3(4) = -5k - 10$. Então $-5k - 10 = 0$, ou $k = -2$.

(b) $u \cdot v = 12 - 3k - 12 + 7 + 10k = 7k + 7$. Então $7k + 7 = 0$, ou $k = -1$.

1.10 Encontre $\|u\|$, nos casos (a) $u = (3, -12, -4)$, (b) $u = (2, -3, 8, -7)$.

Primeiro calculamos $\|u\|^2 = u \cdot u$, somando o quadrado das entradas do vetor. Então $\|u\| = \sqrt{\|u\|^2}$.

(a) $\|u\|^2 = (3)^2 + (-12)^2 + (-4)^2 = 9 + 144 + 16 = 169$. Então $\|u\| = \sqrt{169} = 13$.

(b) $\|u\|^2 = 4 + 9 + 64 + 49 = 126$. Então $\|u\| = \sqrt{126}$.

1.11 Lembre que *normalizar* um vetor não nulo significa encontrar o único vetor unitário $\hat{v}$ com mesma direção e sentido que v, dado por

$$\hat{v} = \frac{1}{\|v\|}\, v$$

Normalize (a) $u = (3, -4)$, (b) $v = (4, -2, -3, 8)$, (c) $w = (\frac{1}{2}, \frac{2}{3}, -\frac{1}{4})$.

(a) Primeiro calculamos $\|u\| = \sqrt{9+16} = \sqrt{25} = 5$. Então dividimos cada entrada de u por 5, obtendo $\hat{u} = (\frac{3}{5}, -\frac{4}{5})$.

(b) Temos $\|v\| = \sqrt{16+4+9+64} = \sqrt{93}$. Logo,

$$\hat{v} = \left(\frac{4}{\sqrt{93}}, \frac{-2}{\sqrt{93}}, \frac{-3}{\sqrt{93}}, \frac{8}{\sqrt{93}}\right)$$

(c) Observe que w, e qualquer múltiplo positivo de w, terão a mesma normalização. Portanto, primeiro multiplicamos w por 12 para eliminar as frações, ou seja, primeiro calculamos $w' = 12w = (6, 8, -3)$. Então

$$\|w'\| = \sqrt{36+64+9} = \sqrt{109} \quad \text{e} \quad \hat{w} = \widehat{w'} = \left(\frac{6}{\sqrt{109}}, \frac{8}{\sqrt{109}}, \frac{-3}{\sqrt{109}}\right)$$

1.12 Sejam $u = (1, -3, 4)$ e $v = (3, 4, 7)$. Encontre

(a) θ, onde θ é o ângulo entre u e v;

(b) a projeção proj(u, v) de u sobre v;

(c) a distância $d(u, v)$ entre u e v.

Primeiro calculamos $u \cdot v = 3 - 12 + 28 = 19$, $\|u\|^2 = 1 + 9 + 16 = 26$, $\|v\|^2 = 9 + 16 + 49 = 74$. Então

(a) $\cos\theta = \dfrac{u \cdot v}{\|u\|\|v\|} = \dfrac{19}{\sqrt{26}\sqrt{74}}$,

(b) $\operatorname{proj}(u, v) = \dfrac{u \cdot v}{\|v\|^2}\, v = \dfrac{19}{74}(3, 4, 7) = \left(\dfrac{57}{74}, \dfrac{76}{74}, \dfrac{133}{74}\right) = \left(\dfrac{57}{74}, \dfrac{38}{37}, \dfrac{133}{74}\right)$,

(c) $d(u, v) = \|u - v\| = \|(-2, -7 - 3)\| = \sqrt{4 + 49 + 9} = \sqrt{62}$.

1.13 Demonstre o Teorema 1.2. Dados quaisquer u, v e w em $\mathbf{R}^n$ e r em $\mathbf{R}$,

(i) $(u + v) \cdot w = u \cdot w + v \cdot w$, (ii) $(ru) \cdot v = r(u \cdot v)$, (iii) $u \cdot v = v \cdot u$,

(iv) $u \cdot u \geq 0$ e $u \cdot u = 0$ se, e só se, $u = 0$.

Sejam $u = (u_1, u_2, ..., u_n)$, $v = (v_1, v_2, ..., v_n)$ e $w = (w_1, w_2, ..., w_n)$.

(i) Como $u + v = (u_1 + v_1,\ u_2 + v_2,\ \ldots,\ u_n + v_n)$,

$$\begin{aligned}(u + v) \cdot w &= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \cdots + (u_n + v_n)w_n \\ &= u_1w_1 + v_1w_1 + u_2w_2 + \cdots + u_nw_n + v_nw_n \\ &= (u_1w_1 + u_2w_2 + \cdots + u_nw_n) + (v_1w_1 + v_2w_2 + \cdots + v_nw_n) \\ &= u \cdot w + v \cdot w\end{aligned}$$

(ii) Como $ru = (ru_1, ru_2, \ldots, ru_n)$,

$$(ru) \cdot v = ru_1v_1 + ru_2v_2 + \cdots + ru_nv_n = r(u_1v_1 + u_2v_2 + \cdots + u_nv_n) = r(u \cdot v)$$

(iii) $u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n = v_1u_1 + v_2u_2 + \cdots + v_nu_n = v \cdot u$

(iv) Como u_k^2, é não negativo, para cada k, e como a soma de números não negativos é não negativa,

$$u \cdot u = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$$

Além disso, $u \cdot u = 0$ se, e só se, $u_k = 0$, para cada k, ou seja, se, e só se, $u = 0$.

1.14 Demonstre o Teorema 1.3 (Schwarz). $|u \cdot v| \leq \|u\|\|v\|$.

Dado qualquer número real t, o Teorema 1.2 garante que

$$0 \leq (tu + v) \cdot (tu + v) = t^2(u \cdot u) + 2t(u \cdot v) + (v \cdot v) = \|u\|^2 t^2 + 2(u \cdot v)t + \|v\|^2$$

Sejam $a = \|u\|^2,\ b = 2(u \cdot v)$ e $c = \|v\|^2$. Então, para cada valor de t, temos $at^2 + bt + c \geq 0$. Isso significa que esse polinômio de segundo grau não pode ter duas raízes reais. Disso decorre que o discriminante $D = b^2 - 4ac \leq 0$ ou, equivalentemente, que $b^2 \leq 4ac$. Assim,

$$4(u \cdot v)^2 \leq 4\|u\|^2\|v\|^2$$

Resta dividir por 4 para obter nosso resultado.

1.15 Demonstre o Teorema 1.4 (Minkowski). $\|u + v\| \leq \|u\| + \|v\|$.

Pela desigualdade de Schwarz e outras propriedades do produto escalar, temos

$$\|u + v\|^2 = (u + v) \cdot (u + v) = (u \cdot u) + 2(u \cdot v) + (v \cdot v) \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

Tirando a raiz quadrada de ambos lados, obtemos a desigualdade procurada.

Pontos, retas e hiperplanos em $\mathbf{R}^n$

Nesta parte vamos distinguir entre uma ênupla $P(a_1, a_2, ..., a_n)$, considerada como sendo um ponto de $\mathbf{R}^n$, e uma ênupla $u = [c_1, c_2, ..., c_n]$, considerada como sendo um vetor (segmento orientado) desde a origem O até o ponto $C(c_1, c_2, ..., c_n)$.

1.16 Encontre o vetor u identificado com o segmento de reta orientado $\overrightarrow{PQ}$ para os pontos
(a) $P(1, -2, 4)$ e $Q(6, 1, -5)$ em $\mathbf{R}^3$, (b) $P(2, 3, -6, 5)$ e $Q(7, 1, 4, -8)$ em $\mathbf{R}^4$.

(a) $u = \overrightarrow{PQ} = Q - P = [6 - 1,\ 1 - (-2),\ -5 - 4] = [5, 3, -9]$

(b) $u = \overrightarrow{PQ} = Q - P = [7 - 2,\ 1 - 3,\ 4 + 6,\ -8 - 5] = [5, -2, 10, -13]$

1.17 Encontre uma equação do hiperplano H de $\mathbf{R}^4$ que passa por $P(3, -4, 1, -2)$ e é normal a $u = [2, 5, -6, -3]$.

Os coeficientes das incógnitas na equação de H são os componentes do vetor normal u. Assim, uma equação de H é da forma $2x_1 + 5x_2 - 6x_3 - 3x_4 = k$. Substituindo P nessa equação, obtemos $k = 1$. Então uma equação de H é $2x_1 + 5x_2 - 6x_3 - 3x_4 = -26$.

1.18 Encontre uma equação do plano H de $\mathbf{R}^3$ que contém o ponto $P(1, -3, -4)$ e é paralelo ao plano H' determinado pela equação $3x - 6y + 5z = 2$.

Os planos H e H' são paralelos se, e só se, suas direções normais são paralelas, com mesmo sentido ou sentidos opostos. Logo, uma equação de H é do tipo $3x - 6y + 5z = k$. Substituindo P nessa equação, obtemos $k = 1$. Então, uma equação de H é $3x - 6y + 5z = 1$.

1.19 Encontre uma representação paramétrica da reta L de $\mathbf{R}^4$ que passa por $P(4, -2, 3, 1)$ na direção de $u = [2, 5, -7, 8]$.

Aqui, L consiste nos pontos $X(x_k)$ que satisfazem

$$X = P + tu, \quad \text{ou} \quad x_i = a_k t + b_k, \quad \text{ou} \quad L(t) = (a_k t + b_k),$$

em que o parâmetro t percorre todos os reais. Então, obtemos

$$x_1 = 4 + 2t,\ x_2 = -2 + 2t,\ x_3 = 3 - 7t,\ x_4 = 1 + 8t \quad \text{ou} \quad L(t) = (4 + 2t,\ -2 + 2t,\ 3 - 7t,\ 1 + 8t)$$

1.20 Seja C a curva dada por $F(t) = (t^2, 3t - 2, t^3, t^2 + 5)$ de $\mathbf{R}^4$, com $0 \leq t \leq 4$.

(a) Encontre o ponto P de C que corresponde a $t = 2$.

(b) Encontre o ponto inicial Q de C e seu ponto final Q'.

(c) Encontre o vetor tangente unitário $\mathbf{T}$ à curva C com $t = 2$.

(a) Substituindo $t = 2$ em $F(t)$, obtemos $P = F(2) = (4, 4, 8, 9)$.

(b) O parâmetro t varia de $t = 0$ até $t = 4$. Portanto, $Q = F(0) = (0, -2, 0, 5)$ e $Q' = F(4) = (16, 10, 64, 21)$.

(c) Tomando a derivada de $F(t)$, ou seja, de cada componente de $F(t)$, obtemos um vetor V que é tangente à curva.

$$V(t) = \frac{dF(t)}{dt} = [2t, 3, 3t^2, 2t]$$

Agora, calculamos V com $t = 2$, ou seja, substituímos $t = 2$ na equação de $V(t)$ para obter $V = V(2) = [4, 3, 12, 4]$. Em seguida, normalizamos V para obter o vetor tangente unitário $\mathbf{T}$ requerido. Temos

$$\|V\| = \sqrt{16 + 9 + 144 + 16} = \sqrt{185} \qquad \text{e} \qquad \mathbf{T} = \left[\frac{4}{\sqrt{185}}, \frac{3}{\sqrt{185}}, \frac{12}{\sqrt{185}}, \frac{4}{\sqrt{185}}\right]$$

Vetores espaciais (de R^3), notação ijk, produto vetorial

1.21 Sejam $u = 2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$, $v = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ e $w = \mathbf{i} + 5\mathbf{j} + 3\mathbf{k}$. Encontre

(a) $u + v$, (b) $2u - 3v + 4w$, (c) $u \cdot v$ e $u \cdot w$, (d) $\|u\|$ e $\|v\|$.

Tratamos os coeficientes de $\mathbf{i}$, $\mathbf{j}$ e $\mathbf{k}$ como se fossem componentes de um vetor de $\mathbf{R}^3$.

(a) Somamos os coeficientes correspondentes para obter $u + v = 5\mathbf{i} - 2\mathbf{j} - 2\mathbf{k}$.

(b) Primeiro multiplicamos por escalar e depois somamos os vetores.

$$\begin{aligned} 2u - 3v + 4w &= (4\mathbf{i} - 6\mathbf{j} + 8\mathbf{k}) + (-9\mathbf{i} + 3\mathbf{j} + 6\mathbf{k}) + (4\mathbf{i} + 20\mathbf{j} + 12\mathbf{k}) \\ &= -\mathbf{i} + 17\mathbf{j} + 26\mathbf{k} \end{aligned}$$

(c) Multiplicamos os coeficientes correspondentes e depois somamos.

$$u \cdot v = 6 - 3 - 8 = -5 \qquad \text{e} \qquad u \cdot w = 2 - 15 + 12 = -1$$

(d) A norma é a raiz quadrada da soma dos quadrados dos coeficientes.

$$\|u\| = \sqrt{4 + 9 + 16} = \sqrt{29} \qquad \text{e} \qquad \|v\| = \sqrt{9 + 1 + 4} = \sqrt{14}$$

1.22 Encontre equações paramétricas da reta L que

(a) passa pelos pontos $P(1, 3, 2)$ e $Q(2, 5, -6)$;

(b) contém o ponto $P(1, -2, 4)$ e é perpendicular ao plano H de equação $3x + 5y + 7z = 15$.

(a) Inicialmente obtemos $v = \overrightarrow{PQ} = Q - P = [1, 2, -8] = \mathbf{i} + 2\mathbf{j} - 8\mathbf{k}$. Então

$$L(t) = (t + 1,\ 2t + 3,\ -8t + 2) = (t + 1)\mathbf{i} + (2t + 3)\mathbf{j} + (-8t + 2)\mathbf{k}$$

(b) Como L é perpendicular a H, a reta L tem a mesma direção do vetor normal $\mathbf{N} = 3\mathbf{i} + 5\mathbf{j} + 7\mathbf{k}$ ao plano H. Assim,

$$L(t) = (3t + 1,\ 5t - 2,\ 7t + 4) = (3t + 1)\mathbf{i} + (5t - 2)\mathbf{j} + (7t + 4)\mathbf{k}$$

1.23 Seja S a superfície de equação $xy^2 + 2yz = 16$ em $\mathbf{R}^3$.

(a) Encontre o vetor normal $\mathbf{N}(x, y, z)$ à superfície S.

(b) Encontre o plano tangente H à superfície S no ponto $P(1, 2, 3)$.

(a) A fórmula do vetor normal a uma superfície dada por $F(x, y, z) = 0$ é

$$\mathbf{N}(x, y, z) = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k}$$

onde F_x, F_y e F_z são as derivadas parciais de F. Usando $F(x, y, z) = xy^2 + 2yz - 16$, obtemos

$$F_x = y^2, \qquad F_y = 2xy + 2z, \qquad F_z = 2y$$

Assim, $\mathbf{N}(x, y, z) = y^2\mathbf{i} + (2xy + 2z)\mathbf{j} + 2y\mathbf{k}$.

(b) O vetor normal à superfície S no ponto P é

$$\mathbf{N}(P) = \mathbf{N}(1, 2, 3) = 4\mathbf{i} + 10\mathbf{j} + 4\mathbf{k}$$

Logo, $\mathbf{N} = 2\mathbf{i} + 5\mathbf{j} + 2\mathbf{k}$ também é normal a S em P. Portanto, uma equação de H tem a forma $2x + 5y + 2z = c$. Substituindo P nessa equação, obtemos $c = 18$. Assim, o plano tangente H a S em P é dado por $2x + 5y + 2z = 18$.

1.24 Calcule os determinantes e os simétricos de determinantes de ordem dois seguintes.

(a) (i) $\begin{vmatrix} 3 & 4 \\ 5 & 9 \end{vmatrix}$, (ii) $\begin{vmatrix} 2 & -1 \\ 4 & 3 \end{vmatrix}$, (iii) $\begin{vmatrix} 4 & -5 \\ 3 & -2 \end{vmatrix}$

(b) (i) $-\begin{vmatrix} 3 & 6 \\ 4 & 2 \end{vmatrix}$, (ii) $-\begin{vmatrix} 7 & -5 \\ 3 & 2 \end{vmatrix}$, (iii) $-\begin{vmatrix} 4 & -1 \\ 8 & -3 \end{vmatrix}$

Usamos $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ e $-\begin{vmatrix} a & b \\ c & d \end{vmatrix} = bc - ad$. Assim,

(a) (i) $27 - 20 = 7$, (ii) $6 + 4 = 10$, (iii) $-8 + 15 = 7$.

(b) (i) $24 - 6 = 18$, (ii) $-15 - 14 = -29$, (iii) $-8 + 12 = 4$.

1.25 Sejam $u = 2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$, $v = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ e $w = \mathbf{i} + 5\mathbf{j} + 3\mathbf{k}$. Encontre (a) $u \times v$, (b) $u \times w$

(a) Usamos $\begin{bmatrix} 2 & -3 & 4 \\ 3 & 1 & -2 \end{bmatrix}$ para obter $u \times v = (6 - 4)\mathbf{i} + (12 + 4)\mathbf{j} + (2 + 9)\mathbf{k} = 2\mathbf{i} + 16\mathbf{j} + 11\mathbf{k}$.

(b) Usamos $\begin{bmatrix} 2 & -3 & 4 \\ 1 & 5 & 3 \end{bmatrix}$ para obter $u \times w = (-9 - 20)\mathbf{i} + (4 - 6)\mathbf{j} + (10 + 3)\mathbf{k} = -29\mathbf{i} - 2\mathbf{j} + 13\mathbf{k}$.

1.26 Encontre $u \times v$, sendo (a) $u = (1, 2, 3)$, $v = (4, 5, 6)$; (b) $u = (-4, 7, 3)$, $v = (6, -5, 2)$.

(a) Usamos $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ para obter $u \times v = [12 - 15,\ 12 - 6,\ 5 - 8] = [-3, 6, -3]$.

(b) Usamos $\begin{bmatrix} -4 & 7 & 3 \\ 6 & -5 & 2 \end{bmatrix}$ para obter $u \times v = [14 + 15,\ 18 + 8,\ 20 - 42] = [29, 26, -22]$.

1.27 Encontre um vetor unitário u ortogonal a $v = [1, 3, 4]$ e $w = [2, -6, -5]$.

Primeiro calculamos $v \times w$, que é ortogonal a v e a w.

A tabela $\begin{bmatrix} 1 & 3 & 4 \\ 2 & -6 & -5 \end{bmatrix}$ nos dá $v \times w = [-15 + 24,\ 8 + 5,\ -6 - 61] = [9, 13, -12]$.

Agora normalizamos $v \times w$ para obter $u = [9/\sqrt{394},\ 13/\sqrt{394},\ -12/\sqrt{394}]$.

1.28 Dados $u = (a_1, a_2, a_3)$ e $v = (b_1, b_2, b_3)$, temos $u \times v = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$. Prove que

(a) $u \times v$ é ortogonal a u e a v [Teorema 1.5(a)].

(b) $\|u \times v\|^2 = (u \cdot u)(v \cdot v) - (u \cdot v)^2$ (Identidade de Lagrange).

(a) Temos

$$\begin{aligned} u\cdot(u\times v) &= a_1(a_2b_3-a_3b_2)+a_2(a_3b_1-a_1b_3)+a_3(a_1b_2-a_2b_1) \\ &= a_1a_2b_3-a_1a_3b_2+a_2a_3b_1-a_1a_2b_3+a_1a_3b_2-a_2a_3b_1=0 \end{aligned}$$

Assim, $u\times v$ é ortogonal a u. Analogamente, $u\times v$ é ortogonal a v.

(b) Temos

$$\|u\times v\|^2 = (a_2b_3-a_3b_2)^2+(a_3b_1-a_1b_3)^2+(a_1b_2-a_2b_1)^2 \quad (1)$$
$$(u\cdot u)(v\cdot v)-(u\cdot v)^2 = (a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2 \quad (2)$$

Expandindo o lado direito das igualdades (1) e (2), estabelecemos a identidade.

Números complexos, vetores de C^n

1.29 Suponha que $z = 5+3i$ e $w = 2-4i$. Encontre (a) $z+w$, (b) $z-w$, (c) zw.

Usamos as regras algébricas usuais, junto com $i^2 = -1$, para obter os resultados no formato padrão $a+bi$.

(a) $z+w=(5+3i)+(2-4i)=7-i$

(b) $z-w=(5+3i)-(2-4i)=5+3i-2+4i=3+7i$

(c) $zw=(5+3i)(2-4i)=10-14i-12i^2=10-14i+12=22-14i$

1.30 Simplifique (a) $(5+3i)(2-7i)$, (b) $(4-3i)^2$, (c) $(1+2i)^3$.

(a) $(5+3i)(2-7i)=10+6i-35i-21i^2=31-29i$

(b) $(4-3i)^2=16-24i+9i^2=7-24i$

(c) $(1+2i)^3=1+6i+12i^2+8i^3=1+6i-12-8i=-11-2i$

1.31 Simplifique (a) i^0, i^3, i^4, (b) i^5, i^6, i^7, i^8, (c) i^{39}, i^{174}, i^{252}, i^{317}.

(a) $i^0=1,\quad i^3=i^2(i)=(-1)(i)=-i,\quad i^4=(i^2)(i^2)=(-1)(-1)=1$

(b) $i^5=(i^4)(i)=(1)(i)=i,\quad i^6=(i^4)(i^2)=(1)(i^2)=i^2=-1,\quad i^7=i^3=-i,\quad i^8=i^4=1$

(c) Usando $i^4=1$ e $i^n=i^{4q+r}=(i^4)^q i^r=1^q i^r=i^r$, dividimos o expoente n por 4 para obter o resto r.

$$i^{39}=i^{4(9)+3}=(i^4)^9i^3=1^9i^3=i^3=-i,\qquad i^{174}=i^2=-1,\qquad i^{252}=i^0=1,\qquad i^{317}=i^1=i$$

1.32 Encontre o conjugado de cada um dos números complexos seguintes.

(a) $6+4i$, $7-5i$, $4+i$, $-3-i$, (b) 6, -3, $4i$, $-9i$.

(a) $\overline{6+4i}=6-4i,\ \overline{7-5i}=7+5i,\ \overline{4+i}=4-i,\ \overline{-3-i}=-3+i$

(b) $\bar{6}=6,\ \overline{-3}=-3,\ \overline{4i}=-4i,\ \overline{-9i}=9i$

(Observe que o conjugado de um número real é o próprio número real, mas o conjugado de um número imaginário puro é seu simétrico.)

1.33 Encontre $z\bar{z}$ e $|z|$ para $z = 3+4i$.

Com $z=a+bi$ usamos $z\bar{z}=a^2+b^2$ e $z=\sqrt{z\bar{z}}=\sqrt{a^2+b^2}$. Logo,

$$z\bar{z}=9+16=25,\qquad |z|=\sqrt{25}=5$$

1.34 Simplifique $\dfrac{2-7i}{5+3i}$.

Para simplificar uma fração z/w de números complexos, multiplicamos pelo conjugado $\bar{w}$ do denominador tanto o numerador quanto o denominador.

$$\frac{2-7i}{5+3i}=\frac{(2-7i)(5-3i)}{(5+3i)(5-3i)}=\frac{-11-41i}{34}=-\frac{11}{34}-\frac{41}{34}i$$

1.35 Demonstre que, para quaisquer números complexos $z, w \in \mathbf{C}$, (i) $\overline{z+w} = \bar{z} + \bar{w}$, (ii) $\overline{zw} = \bar{z}\bar{w}$, (iii) $\bar{\bar{z}} = z$.

Suponha que $z = a + bi$ e $w = c + di$, com $a, b, c, d \in \mathbf{R}$.

(i) $$\overline{z+w} = \overline{(a+bi)+(c+di)} = \overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = a+c-bi-di = (a-bi)+(c-di) = \bar{z}+\bar{w}$$

(ii) $$\overline{zw} = \overline{(a+bi)(c+di)} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i = (a-bi)(c-di) = \bar{z}\bar{w}$$

(iii) $$\bar{\bar{z}} = \overline{\overline{a+bi}} = \overline{a-bi} = a-(-b)i = a+bi = z$$

1.36 Demonstre que, para quaisquer números complexos z e w, vale $|zw| = |z|\,|w|$.

Pela parte (ii) do problema precedente,

$$|zw|^2 = (zw)(\overline{zw}) = (zw)(\bar{z}\bar{w}) = (z\bar{z})(w\bar{w}) = |z|^2|w|^2$$

Tirando a raiz quadrada dos dois lados dessa igualdade obtemos nosso resultado.

1.37 Demonstre que, para quaisquer números complexos z e w, vale $z, w \in \mathbf{C}$, $|z+w| \le |z| + |w|$.

Suponha que $z = a + bi$ e $w = c + di$, com $a, b, c, d \in \mathbf{R}$. Considere os vetores $u = (a, b)$ e $v = (c, d)$ de $\mathbf{R}^2$. Observe que

$$|z| = \sqrt{a^2+b^2} = \|u\|, \qquad |w| = \sqrt{c^2+d^2} = \|v\|$$

e

$$|z+w| = |(a+c)+(b+d)i| = \sqrt{(a+c)^2+(b+d)^2} = \|(a+c, b+d)\| = \|u+v\|$$

Pela desigualdade de Minkowski (Problema 1.15), $\|u+v\| \le \|u\| + \|v\|$ e, portanto,

$$|z+w| = \|u+v\| \le \|u\| + \|v\| = |z| + |w|$$

1.38 Encontre os produtos escalares $u \cdot v$ e $v \cdot u$ nos casos seguintes. (a) $u = (1-2i,\ 3+i)$, $v = (4+2i, 5-6i)$, (b) $u = (3-2i,\ 4i,\ 1+6i)$, $v = (5+i,\ 2-3i,\ 7+2i)$.

Não esqueça que no produto escalar aparecem os conjugados do segundo vetor.

$$(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n$$

(a) $$u \cdot v = (1-2i)\overline{(4+2i)} + (3+i)\overline{(5-6i)} = (1-2i)(4-2i) + (3+i)(5+6i) = -10i + 9 + 23i = 9 + 13i$$

$$v \cdot u = (4+2i)\overline{(1-2i)} + (5-6i)\overline{(3+i)} = (4+2i)(1+2i) + (5-6i)(3-i) = 10i + 9 - 23i = 9 - 13i$$

(b) $$u \cdot v = (3-2i)\overline{(5+i)} + (4i)\overline{(2-3i)} + (1+6i)\overline{(7+2i)} = (3-2i)(5-i) + (4i)(2+3i) + (1+6i)(7-2i) = 20 + 35i$$

$$v \cdot u = (5+i)\overline{(3-2i)} + (2-3i)\overline{(4i)} + (7+2i)\overline{(1+6i)} = (5+i)(3+2i) + (2-3i)(-4i) + (7+2i)(1-6i) = 20 - 35i$$

Em ambos casos, $v \cdot u = \overline{u \cdot v}$. Isso é válido em geral, conforme Problema 1.40.

1.39 Sejam $u = (7-2i, 2+5i)$ e $v = (1+i, -3-6i)$. Encontre

(a) $u + v$, (b) $2iu$, (c) $(3-i)v$, (d) $u \cdot v$, (e) $\|u\|$ e $\|v\|$.

(a) $u + v = (7 - 2i + 1 + i,\ 2 + 5i - 3 - 6i) = (8 - i,\ -1 - i)$

(b) $2iu = (14i - 4i^2,\ 4i + 10i^2) = (4 + 14i,\ -10 + 4i)$

(c) $(3-i)v = (3 + 3i - i - i^2,\ -9 - 18i + 3i + 6i^2) = (4 + 2i,\ -15 - 15i)$

(d) $u \cdot v = (7-2i)\overline{(1+i)} + (2+5i)\overline{(-3-6i)}$
$= (7-2i)(1-i) + (2+5i)(-3+6i) \;=\; 5-9i-36-3i \;=\; -31-12i$

(e) $\|u\| = \sqrt{7^2 + (-2)^2 + 2^2 + 5^2} = \sqrt{82}$ e $\|v\| = \sqrt{1^2 + 1^2 + (-3)^2 + (-6)^2} = \sqrt{47}$

1.40 Demonstre que, para quaisquer vetores $u, v \in \mathbf{C}^n$ e escalar $z \in \mathbf{C}$, (i) $u \cdot v = \overline{v \cdot u}$, (ii) $(zu) \cdot v = z(u \cdot v)$, valem (iii) $u \cdot (zv) = \bar{z}(u \cdot v)$.

Suponha que $u = (z_1, z_2, \ldots, z_n)$ e $v = (w_1, w_2, \ldots, w_n)$.

(i) Usando as propriedades da conjugação,

$$\overline{v \cdot u} = \overline{w_1\bar{z}_1 + w_2\bar{z}_2 + \cdots + w_n\bar{z}_n} = \overline{w_1\bar{z}_1} + \overline{w_2\bar{z}_2} + \cdots + \overline{w_n\bar{z}_n}$$
$$= \bar{w}_1 z_1 + \bar{w}_2 z_2 + \cdots + \bar{w}_n z_n = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n = u \cdot v$$

(ii) Como $zu = (zz_1, zz_2, \ldots, zz_n)$,

$$(zu) \cdot v = zz_1\bar{w}_1 + zz_2\bar{w}_2 + \cdots + zz_n\bar{w}_n = z(z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n) = z(u \cdot v)$$

(Compare essa identidade com a do item (ii) do Teorema 1.2 para vetores de $\mathbf{R}^n$.)

(iii) Usando (i) e (ii),

$$u \cdot (zv) = \overline{(zv) \cdot u} = \overline{z(v \cdot u)} = \bar{z}(\overline{v \cdot u}) = \bar{z}(u \cdot v)$$

Problemas Complementares

Vetores de $\mathbf{R}^n$

1.41 Sejam $u = (1, -2, 4)$, $v = (3, 5, 1)$ e $w = (2, 1, -3)$. Encontre

(a) $3u - 2v$; (b) $5u + 3v - 4w$; (c) $u \cdot v$, $u \cdot w$, $v \cdot w$; (d) $\|u\|$, $\|v\|$;

(e) θ, sendo θ o ângulo entre u e v; (f) $d(u, v)$; (g) proj(u, v).

1.42 Refaça o problema precedente com os vetores $u = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$, $w = \begin{bmatrix} 3 \\ -2 \\ 6 \end{bmatrix}$.

1.43 Sejam $u = (2, -5, 4, 6, -3)$ e $v = (5, -2, 1, -7, -4)$. Encontre

(a) $4u - 3v$; (b) $5u + 2v$; (c) $u \cdot v$; (d) $\|u\|$ e $\|v\|$; (e) proj(u, v); (f) $d(u, v)$.

1.44 Normalize cada vetor dado.

(a) $u = (5, -7)$; (b) $v = (1, 2, -2, 4)$; (c) $w = \left(\frac{1}{2}, -\frac{1}{3}, \frac{3}{4}\right)$.

1.45 Sejam $u = (1, 2, -2)$, $v = (3, -12, 4)$ e $r = -3$.

(a) Encontre $\|u\|$, $\|v\|$, $\|u + v\|$ e $\|ru\|$.

(b) Verifique que $\|ru\| = |r|\,\|u\|$ e que $\|u + v\| \leq \|u\| + \|v\|$.

1.46 Encontre x e y tais que

(a) $(x, y + 1) = (y - 2, 6)$; (b) $x(2, y) = y(1, -2)$.

1.47 Encontre x, y e z tais que $(x, y + 1, y + z) = (2x + y, 4, 3z)$.

1.48 Escreva $v = (2, 5)$ como uma combinação linear de u_1 e u_2, sendo

(a) $u_1 = (1, 2)$ e $u_2 = (3, 5)$;

(b) $u_1 = (3, -4)$ e $u_2 = (2, -3)$.

1.49 Escreva $v = \begin{bmatrix} 9 \\ -3 \\ 16 \end{bmatrix}$ como uma combinação linear de $u_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $u_2 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_3 = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$.

1.50 Encontre um escalar k tal que u e v sejam ortogonais, nos casos seguintes.

(a) $u = (3, k, -2)$, $v = (6, -4, -3)$;

(b) $u = (5, k, -4, 2)$, $v = (1, -3, 2, 2k)$;

(c) $u = (1, 7, k + 2, -2)$, $v = (3, k, -3, k)$.

Vetores aplicados, hiperplanos e retas de R^n

1.51 Encontre o vetor v identificado com o segmento de reta orientado $\overrightarrow{PQ}$ para os pontos dados.

(a) $P(2, 3, -7)$ e $Q(1, -6, -5)$ em $\mathbf{R}^3$;

(b) $P(1, -8, -4, 6)$ e $Q(3, -5, 2, -4)$ em $\mathbf{R}^4$.

1.52 Encontre uma equação do hiperplano H de $\mathbf{R}^4$ que

(a) contenha $P(1, 2, -3, 2)$ e seja normal a $u = [2, 3, -5, 6]$;

(b) contenha $P(3, -1, 2, 5)$ e seja paralelo ao hiperplano $2x_1 - 3x_2 + 5x_3 - 7x_4 = 4$.

1.53 Encontre uma representação paramétrica da reta de $\mathbf{R}^4$ que

(a) passe pelos pontos $P(1, 2, 1, 2)$ e $Q(3, -5, 7, -9)$;

(b) passe por $P(1, 1, 3, 3)$ e seja perpendicular ao hiperplano $2x_1 + 4x_2 + 6x_3 - 8x_4 = 5$.

Vetores espaciais (de R^3), notação ijk

1.54 Sejam $u = 3\mathbf{i} - 4\mathbf{j} + 2\mathbf{k}$, $v = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$ e $w = 4\mathbf{i} + 7\mathbf{j} + 2\mathbf{k}$. Encontre

(a) $2u - 3v$; (b) $3u + 4v - 2w$; (c) $u \cdot v$, $u \cdot w$, $v \cdot w$; (d) $\|u\|$, $\|v\|$, $\|w\|$.

1.55 Encontre uma equação do plano H que

(a) tenha uma normal $\mathbf{N} = 3\mathbf{i} - 4\mathbf{j} + 5\mathbf{k}$ e que contenha o ponto $P(1, 2, -3)$;

(b) seja paralelo ao plano $4x + 3y - 2z = 11$ e que contenha o ponto $Q(2, -1, 3)$;

1.56 Encontre uma equação (paramétrica) da reta L que

(a) passe pelo ponto $P(2, 5, -3)$ na direção do vetor $v = 4\mathbf{i} - 5\mathbf{j} + 7\mathbf{k}$;

(b) seja perpendicular ao plano $2x - 3y + 7z = 4$ e que contenha o ponto $Q(1, -5, 7)$.

1.57 Considere a curva C de $\mathbf{R}^3$, com $0 \le t \le 5$, dada por

$$F(t) = t^3\mathbf{i} - t^2\mathbf{j} + (2t - 3)\mathbf{k}$$

(a) Encontre o ponto P de C que corresponde a $t = 2$;

(b) Encontre o ponto inicial Q e o ponto final Q' de C;

(c) Encontre o vetor tangente unitário $\mathbf{T}$ à curva C que corresponde a $t = 2$.

1.58 Um certo objeto B em movimento tem sua posição no instante de tempo t dada por $R(t) = t^2\mathbf{i} + t^3\mathbf{j} + 2t\mathbf{k}$. [Então $V(t) = dR(t)/dt$ e $A(t) = dV(t)/dt$ denotam a velocidade e a aceleração de B, respectivamente.] Quando $t = 1$, obtenha, para o objeto B, sua

(a) posição R; (b) velocidade V; (c) velocidade escalar $\|V\|$ e (d) aceleração A.

1.59 Encontre um vetor normal **N** e o plano tangente H a cada superfície dada no ponto dado.

(a) A superfície $x^2y + 3yz = 20$ e o ponto $P(1,3,2)$.

(b) A superfície $x^2 + 3y^2 - 5z^2 = 160$ e o ponto $P(3, -2, 1)$.

Produto vetorial

1.60 Calcule os determinantes e os simétricos de determinantes de ordem dois seguintes.

(a) $\begin{vmatrix} 2 & 5 \\ 3 & 6 \end{vmatrix}$, $\begin{vmatrix} 3 & -6 \\ 1 & -4 \end{vmatrix}$, $\begin{vmatrix} -4 & -2 \\ 7 & -3 \end{vmatrix}$

(b) $-\begin{vmatrix} 6 & 4 \\ 7 & 5 \end{vmatrix}$, $-\begin{vmatrix} 1 & -3 \\ 2 & 4 \end{vmatrix}$, $-\begin{vmatrix} 8 & -3 \\ -6 & -2 \end{vmatrix}$

1.61 Sejam $u = 3\mathbf{i} - 4\mathbf{j} + 2\mathbf{k}$, $v = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$ e $w = 4\mathbf{i} + 7\mathbf{j} + 2\mathbf{k}$. Encontre

(a) $u \times v$, (b) $u \times w$, (c) $v \times w$.

1.62 Sejam $u = [2, 1, 3]$, $v = [4, -2, 2]$ e $w = [1, 1, 5]$. Encontre

(a) $u \times v$, (b) $u \times w$, (c) $v \times w$.

1.63 Encontre o volume V do paralelepípedo determinado pelos vetores u, v e w dados no

(a) Problema 1.61 (b) Problema 1.62

1.64 Encontre um vetor unitário u que seja ortogonal aos vetores

(a) $v = [1, 2, 3]$ e $w = [1, -1, 2]$;

(b) $v = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ e $w = 4\mathbf{i} - 2\mathbf{j} - \mathbf{k}$.

1.65 Demonstre as propriedades seguintes do produto vetorial, válidas para quaisquer vetores u, v e w de $\mathbf{R}^3$ e escalar r.

(a) $u \times v = -(v \times u)$

(b) $u \times u = 0$ para qualquer vetor u

(c) $(ru) \times v = r(u \times v) = u \times (rv)$

(d) $u \times (v + w) = (u \times v) + (u \times w)$

(e) $(v + w) \times u = (v \times u) + (w \times u)$

(f) $(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$

Números complexos

1.66 Simplifique:

(a) $(4 - 7i)(9 + 2i)$; (b) $(3 - 5i)^2$; (c) $\dfrac{1}{4 - 7i}$; (d) $\dfrac{9 + 2i}{3 - 5i}$; (e) $(1 - i)^3$.

1.67 Simplifique: (a) $\dfrac{1}{2i}$; (b) $\dfrac{2 + 3i}{7 - 3i}$; (c) i^{15}, i^{25}, i^{34}; (d) $\left(\dfrac{1}{3 - i}\right)^2$.

1.68 Sejam $z = 2 - 5i$ e $w = 7 + 3i$. Encontre: (a) $v + w$; (b) zw; (c) z/w; (d) $\bar{z}, \bar{w}$; (e) $|z|$, $|w|$.

1.69 Mostre que para números complexos z e w quaisquer valem

(a) $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$, (b) $\operatorname{Im} z = \frac{1}{2}(z - \bar{z})$, (c) $zw = 0$ se, e só se, $z = 0$ ou $w = 0$.

Vetores em C^n

1.70 Sejam $u = (1 + 7i, 2 - 6i)$ e $v = (5 - 2i, 3 - 4i)$. Encontre:

(a) $u + v$ (b) $(3 + i)u$ (c) $2iu + (4 + 7i)v$ (d) $u \cdot v$ (e) $\|u\|$ e $\|v\|$.

1.71 Demonstre que, para quaisquer vetores u, v, w em $\mathbf{C}^n$ valem:

(a) $(u+v)\cdot w = u\cdot w + v\cdot w$, (b) $w\cdot(u+v) = w\cdot u + w\cdot v$.

1.72 Demonstre que a norma de $\mathbf{C}^n$ satisfaz as propriedades seguintes.

[N_1] Dado qualquer vetor u, $\|u\| \geq 0$ e $\|u\| = 0$ se, e só se, $u = 0$.

[N_2] Dados quaisquer vetor u e complexo z, $\|zu\| = |z|\,\|u\|$.

[N_3] Para quaisquer vetores u e v, $\|u+v\| \leq \|u\| + \|v\|$.

Respostas dos Problemas Complementares

1.41 (a) $(-3,-16,10)$; (b) $(6,1,35)$; (c) $-3, -12, 8$; (d) $\sqrt{21}, \sqrt{35}, \sqrt{14}$;
(e) $-3/\sqrt{21}\sqrt{35}$; (f) $\sqrt{62}$; (g) $-\frac{3}{35}(3,5,1) = (-\frac{9}{35}, -\frac{15}{35}, -\frac{3}{35})$

1.42 (Vetores coluna) (a) $(-1,7,-22)$; (b) $(-1,26,-29)$; (c) $-15,-27,34$;
(d) $\sqrt{26}, \sqrt{30}$; (e) $-15/(\sqrt{26}\sqrt{30})$; (f) $\sqrt{86}$; (g) $-\frac{15}{30}v = (-1, -\frac{1}{2}, -\frac{5}{2})$

1.43 (a) $(-13,-14,13,45,0)$; (b) $(20,-29,22,16,-23)$; (c) -6; (d) $\sqrt{90}, \sqrt{95}$;
(e) $-\frac{6}{95}v$; (f) $\sqrt{197}$

1.44 (a) $(5/\sqrt{76},\ 9/\sqrt{76})$; (b) $(\frac{1}{5}, \frac{2}{5}, -\frac{2}{5}, \frac{4}{5})$; (c) $(6/\sqrt{133},\ -4\sqrt{133},\ 9\sqrt{133})$

1.45 (a) $3,\ 13,\ \sqrt{120},\ 9$

1.46 (a) $x=-3,\ y=5$; (b) $x=0,\ y=0$, e $x=1,\ y=2$

1.47 $x=-3$, $y=3$, $z=\frac{3}{2}$

1.48 (a) $v = 5u_1 - u_2$; (b) $v = 16u_1 - 23u_2$

1.49 $v = 3u_1 - u_2 + 2u_3$

1.50 (a) 6; (b) 3; (c) $\frac{3}{2}$

1.51 (a) $v = [-1,-9,2]$; (b) $[2,3,6,-10]$

1.52 (a) $2x_1 + 3x_2 - 5x_3 + 6x_4 = 35$; (b) $2x_1 - 3x_2 + 5x_3 - 7x_4 = -16$

1.53 (a) $[2t+1,\ -7t+2,\ 6t+1,\ -11t+2]$; (b) $[2t+1,\ 4t+1,\ 6t+3,\ -8t+3]$

1.54 (a) $-23\mathbf{j} + 13\mathbf{k}$; (b) $9\mathbf{i} - 6\mathbf{j} - 10\mathbf{k}$; (c) $-20, -12, 37$; (d) $\sqrt{29}, \sqrt{38}, \sqrt{69}$

1.55 (a) $3x - 4y + 5z = -20$; (b) $4x + 3y - 2z = -1$

1.56 (a) $[4t+2,\ -5t+5,\ 7t-3]$; (b) $[2t+1,\ -3t-5,\ 7t+7]$

1.57 (a) $P = F(2) = 8\mathbf{i} - 4\mathbf{j} + \mathbf{k}$; (b) $Q = F(0) = -3\mathbf{k}$, $Q' = F(5) = 125\mathbf{i} - 25\mathbf{j} + 7\mathbf{k}$;
(c) $\mathbf{T} = (6\mathbf{i} - 2\mathbf{j} + \mathbf{k})/\sqrt{41}$

1.58 (a) $\mathbf{i} + \mathbf{j} + 2\mathbf{k}$; (b) $2\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$; (c) $\sqrt{17}$; (d) $2\mathbf{i} + 6\mathbf{j}$

1.59 (a) $\mathbf{N} = 6\mathbf{i} + 7\mathbf{j} + 9\mathbf{k}$, $6x + 7y + 9z = 45$; (b) $\mathbf{N} = 6\mathbf{i} - 12\mathbf{j} - 10\mathbf{k}$, $3x - 6y - 5z = 16$

1.60 (a) $-3, -6, 26$; (b) $-2, -10, 34$

1.61 (a) $2\mathbf{i} + 13\mathbf{j} + 23\mathbf{k}$; (b) $-22\mathbf{i} + 2\mathbf{j} + 37\mathbf{k}$; (c) $31\mathbf{i} - 16\mathbf{j} - 6\mathbf{k}$

1.62 (a) $[8, 8, -8]$; (b) $[2, -7, 1]$; (c) $[-12, -18, 6]$

1.63 (a) 145; (b) 24

1.64 (a) $(7, 1, -3)/\sqrt{59}$; (b) $(5\mathbf{i} + 11\mathbf{j} - 2\mathbf{k})/\sqrt{150}$

1.66 (a) $50 - 55i$; (b) $-16 - 30i$; (c) $\frac{1}{65}(4 + 7i)$; (d) $\frac{1}{2}(1 + 3i)$; (e) $-2 - 2i$

1.67 (a) $-\frac{1}{2}i$; (b) $\frac{1}{58}(5 + 27i)$; (c) $-1, i, -1$; (d) $\frac{1}{50}(4 + 3i)$

1.68 (a) $9 - 2i$; (b) $29 - 29i$; (c) $\frac{1}{61}(-1 - 41i)$; (d) $2 + 5i,\ 7 - 3i$; (e) $\sqrt{29}, \sqrt{58}$

1.69 (a) *Sugestão*: Se $zw = 0$, então $|zw| = |z||w| = |0| = 0$.

1.70 (a) $(6 + 5i,\ 5 - 10i)$; (b) $(-4 + 22i,\ 12 - 16i)$; (c) $(-8 - 41i,\ -4 - 33i)$;
(d) $12 + 2i$; (e) $\sqrt{90}, \sqrt{54}$

Capítulo 2

Álgebra de Matrizes

2.1 INTRODUÇÃO

Neste capítulo estudamos as matrizes e suas operações algébricas. Essas matrizes podem ser vistas como tabelas retangulares de elementos, em que cada entrada depende de dois índices (diferente, portanto, dos vetores, em que cada entrada depende de apenas um índice). Os sistemas de equações lineares e suas soluções (Capítulo 3) podem ser entendidos eficientemente por meio da linguagem das matrizes. Além disso, alguns conceitos abstratos que serão introduzidos em capítulos posteriores, como "mudança de base", "transformações lineares" e "formas quadráticas" podem ser representados por essas matrizes (tabelas retangulares). Por outro lado, o tratamento abstrato da Álgebra Linear que veremos mais adiante nos dará uma nova maneira de entender a estrutura dessas matrizes.

As entradas de nossas matrizes provêm de algum corpo arbitrário K, que consideramos fixado. Os elementos de K são ditos *números*, ou *escalares*. Nada será perdido se o leitor simplesmente supor que K é o corpo $\mathbf{R}$ dos números reais.

2.2 MATRIZES

Uma *matriz A sobre um corpo K* ou, simplesmente, uma matriz A (quando K estiver subentendido) é uma tabela retangular de escalares, costumeiramente apresentada no formato seguinte.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

As *linhas* de uma tal matriz A são as m listas horizontais de escalares dadas por

$$(a_{11}, a_{12}, \dots, a_{1n}), \quad (a_{21}, a_{22}, \dots, a_{2n}), \quad \dots, \quad (a_{m1}, a_{m2}, \dots, a_{mn})$$

e as *colunas* de A são as n listas verticais de escalares dadas por

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix}, \quad \begin{bmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix}$$

Observe que o elemento a_{ij}, denominado *ij-ésima entrada*, ou *elemento*, aparece na linha i e na coluna j. Em geral, denotamos uma tal matriz simplesmente escrevendo $A = [a_{ij}]$.

Dizemos que uma matriz com m linhas e n colunas é uma *matriz m por n*, que escrevemos $m \times n$. O par de números m e n é dito o *tamanho* da matriz. Duas matrizes A e B são iguais, e escrevemos $A = B$, se ambas tiverem o mesmo tamanho e se as entradas correspondentes forem iguais. Assim, a igualdade de duas matrizes $m \times n$ é equivalente a um sistema de mn igualdades, uma para cada par de entradas correspondentes.

Uma matriz com apenas uma linha é denominada *matriz linha*, ou *vetor linha*, e uma matriz com apenas uma coluna é denominada *matriz coluna*, ou *vetor coluna*. Dizemos que uma matriz que tem todas as entradas nulas é uma *matriz nula* ou *matriz zero*, sendo denotada por 0.

As matrizes com todas as entradas dadas por números reais são ditas *matrizes reais*, ou *matrizes sobre* **R**. Analogamente, as matrizes com todas as entradas dadas por números complexos são ditas *matrizes complexas*, ou *matrizes sobre* **C**. Neste texto, tratamos praticamente só com matrizes reais e complexas.

Exemplo 2.1

(a) A tabela retangular $A = \begin{bmatrix} 1 & -4 & 5 \\ 0 & 3 & -2 \end{bmatrix}$ é uma matriz 2×3. Suas linhas são (1, –4, 5) e (0, 3, –2) e suas colunas são

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -4 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

(b) a matriz nula 2×4 é dada por $0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(c) Encontre x, y, z, t tais que

$$\begin{bmatrix} x+y & 2z+t \\ x-y & z-t \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 1 & 5 \end{bmatrix}$$

Por definição de igualdade de matrizes, as quatro entradas correspondente devem ser iguais. Assim,

$$x + y = 3, \qquad x - y = 1, \qquad 2z + t = 7, \qquad z - t = 5$$

Resolvendo esse sistema de equações, obtemos $x = 2$, $y = 1$, $z = 4$, $t = -1$.

2.3 SOMA DE MATRIZES E MULTIPLICAÇÃO POR ESCALAR

Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ duas matrizes de mesmo tamanho, digamos, $m \times n$. A *soma* de A e B, denotada por $A + B$, é a matriz obtida pela soma de elementos correspondentes de A e B, ou seja,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

O *múltiplo* da matriz A pelo escalar k, denotado por $k \cdot A$ ou, simplesmente, kA, é a matriz obtida pelo produto de cada elemento de A por k, ou seja,

$$kA = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}$$

Observe que ambas, $A + B$ e kA, são matrizes $m \times n$. Também definimos

$$-A = (-1)A \qquad \text{e} \qquad A - B = A + (-B)$$

Dizemos que $-A$ é a matriz *simétrica* de A e que $A - B$ é a matriz *diferença* de A e B. Não se define a soma de matrizes de tamanhos distintos.

Exemplo 2.2 Sejam dadas $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 4 & 5 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & 6 & 8 \\ 1 & -3 & -7 \end{bmatrix}$. Então

$$A + B = \begin{bmatrix} 1+4 & -2+6 & 3+8 \\ 0+1 & 4+(-3) & 5+(-7) \end{bmatrix} = \begin{bmatrix} 5 & 4 & 11 \\ 1 & 1 & -2 \end{bmatrix}$$

$$3A = \begin{bmatrix} 3(1) & 3(-2) & 3(3) \\ 3(0) & 3(4) & 3(5) \end{bmatrix} = \begin{bmatrix} 3 & -6 & 9 \\ 0 & 12 & 15 \end{bmatrix}$$

$$2A - 3B = \begin{bmatrix} 2 & -4 & 6 \\ 0 & 8 & 10 \end{bmatrix} + \begin{bmatrix} -12 & -18 & -24 \\ -3 & 9 & 21 \end{bmatrix} = \begin{bmatrix} -10 & -22 & -18 \\ -3 & 17 & 31 \end{bmatrix}$$

Dizemos que a matriz $2A - 3B$ é uma *combinação linear* de A e B.

Em seguida apresentamos as propriedades básicas da soma e da multiplicação por escalar de matrizes.

Teorema 2.1 Sejam A, B e C matrizes quaisquer (de mesmo tamanho) e k e k' escalares quaisquer. Então

(i) $(A + B) + C = A + (B + C)$, (v) $k(A + B) = kA + kB$,

(ii) $A + 0 = 0 + A = A$, (vi) $(k + k')A = kA + k'A$,

(iii) $A + (-A) = (-A) + A = 0$, (vii) $(kk')A = k(k'A)$,

(iv) $A + B = B + A$, (viii) $1 \cdot A = A$.

Observe que, nos itens (ii) e (iii) do teorema, o 0 se refere à matriz nula. Também, por (i) e (iv), podemos escrever qualquer soma finita de matrizes

$$A_1 + A_2 + \cdots + A_n$$

sem utilizar parênteses, sendo que essa soma não depende da ordem das matrizes. Além disso, usando (vi) e (viii), também temos

$$A + A = 2A, \qquad A + A + A = 3A, \qquad ...$$

e assim por diante.

A demonstração do Teorema 2.1 se reduz a mostrar que são iguais os ij-ésimos elementos de cada lado das equações matriciais. (Ver Problema 2.3.)

Observe a semelhança que ocorre entre o Teorema 2.1 para matrizes e o Teorema 1.1 para vetores. Na verdade, essas operações matriciais podem ser vistas como generalizações das operações vetoriais correspondentes.

2.4 O SÍMBOLO DE SOMATÓRIO

Antes de definir a multiplicação de matrizes, convém introduzir o *símbolo de somatório* Σ (a letra grega maiúscula sigma).

Suponha que $f(k)$ seja uma expressão algébrica a uma variável k. Então a expressão

$$\sum_{k=1}^{n} f(k) \quad \text{ou, equivalentemente,} \quad \textstyle\sum_{k=1}^{n} f(k)$$

tem o significado como segue. Inicialmente, tomamos $k = 1$ em $f(k)$ e obtemos

$$f(1)$$

Em seguida, tomamos $k = 2$ em $f(k)$ e obtemos $f(2)$ que, somado com $f(1)$, fornece

$$f(1) + f(2)$$

Em seguida, tomamos $k = 3$ em $f(k)$ e obtemos $f(3)$ que, somado com a soma precedente, fornece

$$f(1) + f(2) + f(3)$$

Continuamos esse processo até obter a soma

$$f(1) + f(2) + \cdots + f(n)$$

Observe que, a cada passo, aumentamos o valor de k em 1 unidade até alcançarmos n. Dizemos que a letra k é o *índice* do somatório e que 1 e n são, respectivamente, as extremidades *inferior* e *superior* do somatório. No lugar de k frequentemente utilizamos outras letras, como i ou j, para índice.

Também generalizamos nossa definição permitindo que a soma varie desde algum inteiro n_1 qualquer até algum inteiro n_2 qualquer, ou seja, definimos

$$\sum_{k=n_1}^{n_2} f(k) = f(n_1) + f(n_1 + 1) + f(n_1 + 2) + \cdots + f(n_2)$$

Exemplo 2.3

(a) $\sum_{k=1}^{5} x_k = x_1 + x_2 + x_3 + x_4 + x_5$ e $\sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

(b) $\sum_{j=2}^{5} j^2 = 2^2 + 3^2 + 4^2 + 5^2 = 54$ e $\sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$

(c) $\sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} + \cdots + a_{ip} b_{pj}$

2.5 MULTIPLICAÇÃO DE MATRIZES

O *produto* das matrizes A e B, denotado por AB, é um pouco mais complicado. Por isso, começamos com um caso especial.

O produto AB de uma matriz linha $A = [a_i]$ e uma matriz coluna $B = [b_j]$ com o mesmo número de elementos é definido como o escalar (ou a matriz 1×1) obtido pela soma dos produtos das entradas correspondentes, ou seja,

$$AB = [a_1, a_2, \ldots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{k=1}^{n} a_k b_k$$

Enfatizamos que, nesse caso, AB é um escalar (ou, então, uma matriz 1×1). Não definimos o produto AB se a matriz linha A e a matriz coluna B possuírem um número distinto de elementos.

Exemplo 2.4

(a) $[7, -4, 5] \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = 7(3) + (-4)(2) + 5(-1) = 21 - 8 - 5 = 8$

(b) $[6, -1, 8, 3] \begin{bmatrix} 4 \\ -9 \\ -2 \\ 5 \end{bmatrix} = 24 + 9 - 16 + 15 = 32$

Agora, estamos preparados para definir o produto de matrizes mais gerais.

DEFINIÇÃO Sejam $A = [a_{ik}]$ e $B = [b_{kj}]$ duas matrizes tais que o número de colunas de A seja igual ao número de linhas de B, digamos p. Ou seja, supomos que A seja uma matriz $m \times p$ e B uma matriz $p \times n$. Então o *produto* AB de A e B é a matriz $m \times n$ cuja ij-ésima entrada é dada pelo produto da i-ésima linha de A com a j-ésima coluna de B. Assim,

$$\begin{bmatrix} a_{11} & \dots & a_{1p} \\ \cdot & \dots & \cdot \\ a_{i1} & \dots & a_{ip} \\ \cdot & \dots & \cdot \\ a_{m1} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ \cdot & \dots & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \dots & \cdot \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \cdot & \dots & \cdot \\ \cdot & c_{ij} & \cdot \\ \cdot & \dots & \cdot \\ c_{m1} & \dots & c_{mn} \end{bmatrix}$$

onde $\quad c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$

Não definimos o produto AB quando A é uma matriz $m \times p$ e B é uma matriz $q \times n$ com $p \neq q$.

Exemplo 2.5

(a) Encontre AB se $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 0 & -4 \\ 5 & -2 & 6 \end{bmatrix}$.

Como A tem tamanho 2×2 e B 2×3, o produto AB está definido e é uma matriz 2×3. Para obter a primeira linha da matriz produto AB, multiplicamos a primeira linha [1, 3] de A com cada uma das três colunas

$$\begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

de B. Ou seja,

$$AB = \begin{bmatrix} 2+15 & 0-6 & -4+18 \\ & & \end{bmatrix} = \begin{bmatrix} 17 & -6 & 14 \\ & & \end{bmatrix}$$

Para obter a segunda linha de AB, multiplicamos a segunda linha [2, –1] de A com cada coluna de B. Assim,

$$AB = \begin{bmatrix} 17 & -6 & 14 \\ 4-5 & 0+2 & -8-6 \end{bmatrix} = \begin{bmatrix} 17 & -6 & 14 \\ -1 & 2 & -14 \end{bmatrix}$$

(b) Sejam $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 5 & 6 \\ 0 & -2 \end{bmatrix}$. Então

$$AB = \begin{bmatrix} 5+0 & 6-4 \\ 15+0 & 18-8 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 15 & 10 \end{bmatrix} \quad \text{e} \quad BA = \begin{bmatrix} 5+18 & 10+24 \\ 0-6 & 0-8 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ -6 & -8 \end{bmatrix}$$

Esse exemplo mostra que a multiplicação matricial não é comutativa, ou seja, em geral, $AB \neq BA$. No entanto, a multiplicação matricial satisfaz as propriedades seguintes.

Teorema 2.2 Sejam A, B e C matrizes. Sempre que os produtos e somas envolvidos estiverem definidos, valem

(i) $(AB)C = A(BC)$ (associatividade do produto),
(ii) $A(B + C) = AB + AC$ (distributividade à esquerda),
(iii) $(A + B)C = AC + BC$ (distributividade à direita),
(iv) $k(AB) = (kA)B = A(kB)$, onde k é um escalar.

Observe que $A0 = 0$ e $0B = 0$, onde 0 é a matriz nula.

2.6 TRANSPOSTA DE UMA MATRIZ

A *transposta* de uma matriz A, denotada por A^T, é a matriz obtida escrevendo as colunas de A, na mesma ordem, como linhas. Por exemplo,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \qquad \text{e} \qquad [1, -3, -5]^T = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}$$

Em outras palavras, se $A = [a_{ij}]$ é uma matriz $m \times n$, então $A^T = [b_{ij}]$ é a matriz $n \times m$ dada por $b_{ij} = a_{ji}$.

Observe que a transposta de um vetor linha é um vetor coluna. Analogamente, a transposta de um vetor coluna é um vetor linha.

O teorema a seguir enumera as propriedades básicas da transposição.

Teorema 2.3 Sejam A e B matrizes e k um escalar. Então, sempre que os produtos e somas envolvidos estiverem definidos, valem

(i) $(A + B)^T = A^T + B^T$, (iii) $(kA)^T = kA^T$,

(ii) $(A^T)^T = A$, (iv) $(AB)^T = B^T A^T$.

Enfatizamos que, por (iv), a transposta de um produto é o produto das transpostas, só que na ordem inversa.

2.7 MATRIZES QUADRADAS

Uma matriz é dita *quadrada* se tiver o mesmo número de linhas e colunas. Dizemos que uma matriz quadrada $n \times n$ é de *ordem n*.

Vimos que nem sempre é possível somar ou multiplicar duas matrizes quaisquer. No entanto, considerando apenas matrizes quadradas de alguma dada ordem n, esse inconveniente desaparece. Em outras palavras, as operações de adição, multiplicação, multiplicação por escalar e transposição podem sempre ser efetuadas com matrizes quadradas $n \times n$, e o resultado da operação é uma matriz quadrada $n \times n$.

Exemplo 2.6 Duas matrizes quadradas de ordem 3 são as seguintes.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{bmatrix}$$

As seguintes matrizes também são matrizes de ordem 3.

$$A + B = \begin{bmatrix} 3 & -3 & 4 \\ -4 & -1 & -6 \\ 6 & 8 & 3 \end{bmatrix}, \qquad 2A = \begin{bmatrix} 2 & 4 & 6 \\ -8 & -8 & -8 \\ 10 & 12 & 14 \end{bmatrix}, \qquad A^T = \begin{bmatrix} 1 & -4 & 5 \\ 2 & -4 & 6 \\ 3 & -4 & 7 \end{bmatrix}$$

$$AB = \begin{bmatrix} 5 & 7 & -15 \\ -12 & 0 & 20 \\ 17 & 7 & -35 \end{bmatrix}, \qquad BA = \begin{bmatrix} 27 & 30 & 33 \\ -22 & -24 & -26 \\ -27 & -30 & -33 \end{bmatrix}$$

Diagonal e traço

Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n. A *diagonal*, ou *diagonal principal* de A consiste nos elementos com índices iguais, isto é,

$$a_{11}, \quad a_{22}, \quad a_{33}, \quad \ldots, \quad a_{nn}$$

O *traço* de A, denotado por $\mathrm{tr}(A)$, é a soma dos elementos da diagonal principal, a saber,

$$\mathrm{tr}(A) = a_{11} + a_{22} + a_{33} + \cdots + a_{nn}$$

Vale o teorema a seguir.

Teorema 2.4 Sejam A e B matrizes quadradas de mesma ordem e k um escalar. Então valem

(i) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$, (iii) $\mathrm{tr}(A^T) = \mathrm{tr}(A)$,

(ii) $\mathrm{tr}(kA) = k\ \mathrm{tr}(A)$, (iv) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Exemplo 2.7 Sejam A e B as duas matrizes quadradas do Exemplo 2.6. Então

$$\text{diagonal de } A = \{1, -4, 7\} \quad \text{e} \quad \mathrm{tr}(A) = 1 - 4 + 7 = 4$$
$$\text{diagonal de } B = \{2, 3, -4\} \quad \text{e} \quad \mathrm{tr}(B) = 2 + 3 - 4 = 1$$

Além disso,

$$\mathrm{tr}(A + B) = 3 - 1 + 3 = 5, \qquad \mathrm{tr}(2A) = 2 - 8 + 14 = 8, \qquad \mathrm{tr}(A^T) = 1 - 4 + 7 = 4$$
$$\mathrm{tr}(AB) = 5 + 0 - 35 = -30, \qquad \mathrm{tr}(BA) = 27 - 24 - 33 = -30$$

Conforme afirma o Teorema 2.4, temos

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B), \qquad \mathrm{tr}(A^T) = \mathrm{tr}(A), \qquad \mathrm{tr}(2A) = 2\ \mathrm{tr}(A)$$

Além disso, mesmo tendo $AB \neq BA$, os traços desses produtos são iguais.

Matriz identidade, matrizes escalares

A matriz *identidade*, ou matriz *unitária*, de ordem n, denotada por I_n ou, simplesmente, por I, é a matriz quadrada com 1 na diagonal principal e 0 em todas as demais entradas. A matriz identidade I é análoga ao escalar 1, pois, dada qualquer matriz A, temos

$$AI = IA = A$$

Mais geralmente, se B é uma matriz $m \times n$, então $BI_n = I_mB = B$.

Dado qualquer escalar k, dizemos que a matriz kI, com k na diagonal principal e 0 em todas as demais entradas, é a *matriz escalar* correspondente ao escalar k. Observe que

$$(kI)A = k(IA) = kA$$

Ou seja, multiplicar uma matriz A pela matriz escalar kI é equivalente a multiplicar A pelo escalar k.

Exemplo 2.8 As matrizes identidade de ordem 3 e 4 e as matrizes escalares correspondentes a $k = 5$ são as seguintes.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \qquad \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \qquad \begin{bmatrix} 5 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{bmatrix}$$

OBSERVAÇÃO 1 É costume omitir blocos ou padrões de zeros quando não houver dúvidas sobre essas entradas, como fizemos na segunda e quarta matrizes do exemplo.

OBSERVAÇÃO 2 A função *delta de Kronecker* δ_{ij} é definida por

$$\delta_{ij} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

Assim, podemos escrever a matriz identidade como $I = [\delta_{ij}]$.

2.8 POTÊNCIAS DE MATRIZES, POLINÔMIOS MATRICIAIS

Seja A uma matriz quadrada de ordem n com entradas de algum corpo K. As *potências* de A são definidas como segue.

$$A^2 = AA, \qquad A^3 = A^2A, \qquad \ldots, \qquad A^{n+1} = A^nA, \qquad \ldots, \qquad \text{e} \qquad A^0 = I$$

Também definimos polinômios da matriz A. Mais precisamente, dado qualquer polinômio

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

em que os coeficientes a_i são escalares de K, definimos a matriz $f(A)$ como segue.

$$f(A) = a_0I + a_1A + a_2A^2 + \cdots + a_nA^n$$

[Observe que $f(A)$ é obtida de $f(x)$ substituindo a variável x pela matriz A e substituindo o escalar a_0 pela matriz escalar a_0I.] Se $f(A)$ for a matriz nula, dizemos que A é um *zero* ou *raiz* de $f(x)$.

Exemplo 2.9 Se $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$, então

$$A^2 = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} 7 & -6 \\ -9 & 22 \end{bmatrix} \text{ e } A^3 = A^2A = \begin{bmatrix} 7 & -6 \\ -9 & 22 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} -11 & 38 \\ 57 & -106 \end{bmatrix}$$

Sejam $f(x) = 2x^2 - 3x + 5$ e $g(x) = x^2 + 3x - 10$. Então

$$f(A) = 2\begin{bmatrix} 7 & -6 \\ -9 & 22 \end{bmatrix} - 3\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} + 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 16 & -18 \\ -27 & 61 \end{bmatrix}$$

$$g(A) = \begin{bmatrix} 7 & -6 \\ -9 & 22 \end{bmatrix} + 3\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} - 10\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Assim, A é um zero do polinômio $g(x)$.

2.9 MATRIZES INVERTÍVEIS (OU NÃO SINGULARES)

Uma matriz quadrada A é dita *invertível*, ou *não singular*, se existir uma matriz B tal que

$$AB = BA = I$$

onde I é a matriz identidade. Uma tal matriz B é única. Ou seja, se $AB_1 = B_1A = I$ e $AB_2 = B_2A = I$, então

$$B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$$

Dizemos que uma tal matriz B é a *inversa* de A e a denotamos por A^{-1}. Observe que a relação que define a inversa é simétrica; ou seja, se B for a inversa de A, então A será a inversa de B.

Exemplo 2.10 Suponha que $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$. Então

$$AB = \begin{bmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{e} \qquad BA = \begin{bmatrix} 6-5 & 15-15 \\ -2+2 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Assim, B é inversa de A e A é inversa de B.

Sabe-se (Teorema 3.18) que $AB = I$ se, e só se, $BA = I$. Assim, é necessário testar apenas um produto para determinar se duas matrizes são, ou não, inversas. (Ver Problema 2.17.)

Agora suponha que A e B sejam invertíveis. Então AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$. Mais geralmente, se A_1, A_2, ..., A_k forem invertíveis, então seu produto é invertível e a inversa

$$(A_1A_2\ldots A_k)^{-1} = A_k^{-1}\ldots A_2^{-1}A_1^{-1}$$

é o produto das inversas na ordem inversa.

Inversa de uma matriz 2 × 2

Seja A uma matriz 2×2 arbitrária, digamos, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Queremos deduzir uma fórmula para a inversa A^{-1} de A. Mais precisamente, queremos encontrar $2^2 = 4$ escalares, digamos, x_1, y_1, x_2 e y_2 tais que

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} ax_1 + by_1 & ax_2 + by_2 \\ cx_1 + dy_1 & cx_2 + dy_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Igualando as quatro entradas com as entradas correspondentes da matriz identidade, obtemos quatro equações, que podem ser repartidas em dois sistemas 2×2, como segue.

$$\begin{aligned} ax_1 + by_1 &= 1, \qquad & ax_2 + by_2 &= 0 \\ cx_1 + dy_1 &= 0, \qquad & cx_2 + dy_2 &= 1 \end{aligned}$$

Denotemos por $|A| = ab - bc$ o assim chamado *determinante* de A. Supondo que $|A| \neq 0$, podemos resolver em x_1, y_1, x_2 e y_2 de maneira única, obtendo

$$x_1 = \frac{d}{|A|}, \qquad y_1 = \frac{-c}{|A|}, \qquad x_2 = \frac{-b}{|A|}, \qquad y_2 = \frac{a}{|A|}$$

Segue que

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Em outras palavras, quando $|A| \neq 0$, a inversa de uma matriz A de ordem 2 pode ser obtida de A como segue.

(1) Trocamos de lugar os dois elementos da diagonal principal.
(2) Trocamos o sinal dos dois outros elementos.
(3) Multiplicamos a matriz resultante por $1/|A|$ ou, equivalentemente, dividimos cada elemento por $|A|$.

Se $|A| = 0$, então a matriz A não é invertível.

Exemplo 2.11 Encontre a inversa de $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

Primeiro calculamos $|A| = 2(5) - 3(4) = 10 - 12 = -2$. Como $|A| \neq 0$, a matriz A é invertível e

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{5}{2} & \frac{3}{2} \\ 2 & -1 \end{bmatrix}$$

Agora calculamos $|B| = 1(6) - 3(2) = 6 - 6 = 0$. Como $|B| = 0$, a matriz B não tem inversa.

OBSERVAÇÃO A propriedade de uma matriz ser invertível se, e só se, seu determinante for não nulo vale para matrizes quadradas de qualquer ordem. (Ver Capítulo 8.)

Inversa de uma matriz $n \times n$

Suponha que A seja uma matriz quadrada de ordem n qualquer. Encontrar a inversa A^{-1} de A se reduz, também nesse caso, a encontrar a solução de uma coleção de n sistemas $n \times n$ de equações lineares. A solução de tais sistemas e uma maneira eficiente de resolvê-los serão abordadas no Capítulo 3.

2.10 TIPOS ESPECIAIS DE MATRIZES QUADRADAS

Nesta seção descrevemos alguns tipos especiais de matrizes quadradas.

Matrizes diagonais e triangulares

Dizemos que uma matriz quadrada $D = [d_{ij}]$ é *diagonal* se todos os seus elementos fora da diagonal principal forem nulos. Às vezes, denotamos uma tal matriz por

$$D = \text{diag}(d_{11}, d_{22}, ..., d_{nn})$$

em que alguns d_{ii}, ou todos, podem ser nulos. Por exemplo,

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 4 & 0 \\ 0 & -5 \end{bmatrix}, \qquad \begin{bmatrix} 6 & & & \\ & 0 & & \\ & & -9 & \\ & & & 8 \end{bmatrix}$$

são matrizes diagonais que podem ser representadas, respectivamente, por

diag(3, –7, 2), diag(4, –5), diag(6, 0, –9, 8)

(Observe que o padrão de zeros foi omitido na terceira matriz.)

Uma matriz quadrada $A = [a_{ij}]$ é *triangular superior* ou, simplesmente, *triangular* se todas as suas entradas abaixo da diagonal (principal) forem nulas, ou seja, se $a_{ij} = 0$, para $i > j$. Algumas matrizes triangulares superiores arbitrárias de ordens 2, 3 e 4 são dadas a seguir.

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}, \qquad \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ & b_{22} & b_{23} \\ & & b_{33} \end{bmatrix}, \qquad \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ & c_{22} & c_{23} & c_{24} \\ & & c_{33} & c_{34} \\ & & & c_{44} \end{bmatrix}$$

(Assim como nas matrizes diagonais, é costume omitir padrões de zeros.)

Vale o teorema a seguir.

Teorema 2.5 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes $n \times n$ triangulares (superiores). Então

(i) $A + B$, kA e AB são triangulares com respectivas diagonais dadas por
$(a_{11} + b_{11}, \ \ldots, \ a_{nn} + b_{nn})$, $\quad (ka_{11}, \ \ldots, \ ka_{nn})$, $\quad (a_{11}b_{11}, \ \ldots, \ a_{nn}b_{nn})$

(ii) Dado qualquer polinômio $f(x)$, a matriz $f(A)$ é triangular com diagonal

$$(f(a_{11}), f(a_{22}), \ldots, f(a_{nn}))$$

(iii) A é invertível se, e só se, cada elemento diagonal de A for não nulo, ou seja, $a_{ii} \neq 0$; além disso, se existir, a inversa A^{-1} também é triangular.

Uma matriz *triangular inferior* é uma matriz quadrada cujas entradas acima da diagonal são todas nulas. Observamos que o Teorema 2.5 também vale substituindo "triangular" tanto por "triangular inferior" quanto "diagonal".

OBSERVAÇÃO Uma coleção não vazia de matrizes é dita uma *álgebra matricial* se for fechada em relação às operações de adição, multiplicação por escalar e multiplicação matriciais. Claramente, a coleção das matrizes quadradas de alguma ordem fixada constitui uma álgebra de matrizes, bem como a das matrizes escalares, diagonais, triangulares e triangulares inferiores.

Matrizes quadradas reais especiais: simétricas, ortogonais, normais [opcional até o Capítulo 12]

Seja, agora, A uma matriz quadrada com entradas reais, ou seja, uma matriz quadrada real. A relação entre A e sua transposta A^T fornece espécies importantes de matrizes.

(a) Matrizes simétricas

Dizemos que uma matriz A é *simétrica* se $A^T = A$. Equivalentemente, $A = [a_{ij}]$ é simétrica se seus *elementos simétricos* (aqueles que são espelhados pela diagonal) forem iguais, ou seja, se $a_{ij} = a_{ji}$, para cada ij.

Uma matriz A é dita *antissimétrica* se $A^T = -A$ ou, equivalentemente, se $a_{ij} = -a_{ji}$, para cada ij. Claramente, os elementos diagonais de uma tal matriz devem ser, todos, nulos, pois $a_{ii} = -a_{ii}$ implica $a_{ii} = 0$.

(Observe que, se $A^T = A$ ou $A^T = -A$, então A necessariamente é quadrada.)

Exemplo 2.12 Sejam $A = \begin{bmatrix} 2 & -3 & 5 \\ -3 & 6 & 7 \\ 5 & 7 & -8 \end{bmatrix}, B = \begin{bmatrix} 0 & 3 & -4 \\ -3 & 0 & 5 \\ 4 & -5 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(a) Visualmente, os elementos simétricos de A são iguais, ou $A^T = A$. Assim, A é simétrica.
(b) Os elementos diagonais de B são 0 e os elementos simétricos têm sinal oposto, ou $B^T = -B$. Assim, B é antissimétrica.
(c) Como C não é quadrada, não pode ser nem simétrica nem antissimétrica.

(b) Matrizes ortogonais

Uma matriz real é *ortogonal* se $A^T = A^{-1}$ ou, ou seja, se $AA^T = A^TA = I$. Assim, necessariamente, toda matriz ortogonal é quadrada e invertível.

Exemplo 2.13 Seja $A = \begin{bmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{bmatrix}$. Multiplicando A por A^T, obtemos I, ou seja, $AA^T = I$. Isso significa que, também, $A^TA = I$. Assim, $A^T = A^{-1}$, ou seja, A é ortogonal.

Seja, agora, A uma matriz real ortogonal 3×3 com linhas

$$u_1 = (a_1, a_2, a_3), \qquad u_2 = (b_1, b_2, b_3), \qquad u_3 = (c_1, c_2, c_3)$$

Por ser ortogonal, temos $AA^T = I$, ou seja,

$$AA^T = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

Multiplicando A por A^T e igualando cada entrada com a correspondente de I, obtemos as nove equações seguintes.

$$\begin{array}{rrr} a_1^2 + a_2^2 + a_3^2 = 1, & a_1b_1 + a_2b_2 + a_3b_3 = 0, & a_1c_1 + a_2c_2 + a_3c_3 = 0 \\ b_1a_1 + b_2a_2 + b_3a_3 = 0, & b_1^2 + b_2^2 + b_3^2 = 1, & b_1c_1 + b_2c_2 + b_3c_3 = 0 \\ c_1a_1 + c_2a_2 + c_3a_3 = 0, & c_1b_1 + c_2b_2 + c_3b_3 = 0, & c_1^2 + c_2^2 + c_3^2 = 1 \end{array}$$

Vemos que $u_1 \cdot u_1 = 1$, $u_2 \cdot u_2 = 1$, $u_3 \cdot u_3 = 1$ e $u_i \cdot u_j = 0$, com $i \neq j$. Assim, as linhas u_1, u_2, u_3 de A são vetores unitários e dois a dois ortogonais.

Em geral, dizemos que os vetores $u_1, u_2, ..., u_m$ de $\mathbf{R}^n$ constituem um *conjunto ortonormal* se os vetores forem unitários e dois a dois ortogonais, ou seja,

$$u_i \cdot u_j = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

Em outras palavras, $u_i \cdot u_j = \delta_{ij}$, onde δ_{ij} é o delta de Kronecker.

Mostramos que a condição $AA^T = I$ implica que as linhas de A constituem um conjunto ortonormal de vetores. Do mesmo modo, a condição $A^TA = I$ implica que também as colunas de A formam um conjunto ortonormal de vetores. Além disso, como cada passo do argumento é reversível, a recíproca também é verdadeira.

Esses resultados para matrizes 3×3 são verdadeiros em geral. Isto é, vale o teorema a seguir.

Teorema 2.6 Seja A uma matriz real. As afirmações seguintes são equivalentes.

(i) A é ortogonal.
(ii) As linhas de A constituem um conjunto ortonormal.
(iii) As colunas de A constituem um conjunto ortonormal.

Para $n = 2$, temos o resultado seguinte (demonstrado no Problema 2.28).

Teorema 2.7 Seja A uma matriz real e ortogonal 2×2. Então, para algum número real θ, temos

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

(c) Matrizes normais

Uma matriz real A é dita *normal* se comutar com sua transposta A^T, ou seja, se $AA^T = A^TA$. Se A for simétrica, ortogonal ou antissimétrica, então A é normal. Também há outras matrizes normais.

Exemplo 2.14 Seja $A = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix}$. Então

$$AA^T = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix} \quad \text{e} \quad A^TA = \begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix}\begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix}$$

Como $AA^T = A^TA$, a matriz A é normal.

2.11 MATRIZES COMPLEXAS

Seja A uma matriz complexa, isto é, uma matriz com entradas complexas. Na Seção 1.7 vimos que, se $z = a + bi$ for um número complexo, então seu conjugado é $\bar{z} = a - bi$. A *conjugada* de uma matriz complexa A, denotada por $\bar{A}$, é a matriz obtida de A tomando o conjugado de cada entrada de A. Ou seja, se $A = [a_{ij}]$, então $\bar{A} = [b_{ij}]$, com $b_{ij} = \bar{a}_{ij}$. (Isso é denotado escrevendo $\bar{A} = [\bar{a}_{ij}]$.)

As duas operações de transposição e conjugação comutam para qualquer matriz complexa A, e utilizamos a notação especial A^H para a transposta conjugada da matriz A. Ou seja,

$$A^H = (\bar{A})^T = (\overline{A^T})$$

Observe que, se A for real, então $A^H = A^T$. [Muitos livros usam A* no lugar de A^H.]

Exemplo 2.15 Seja $A = \begin{bmatrix} 2+8i & 5-3i & 4-7i \\ 6i & 1-4i & 3+2i \end{bmatrix}$. Então $A^H = \begin{bmatrix} 2-8i & -6i \\ 5+3i & 1+4i \\ 4+7i & 3-2i \end{bmatrix}$.

Matrizes complexas especiais: hermitianas, unitárias, normais [opcional até o Capítulo 12]

Considere uma matriz complexa A. A relação entre A e sua transposta conjugada A^H fornece espécies importantes de matrizes complexas (análogas às espécies de matrizes reais que já vimos).

Dizemos que uma matriz complexa A é *hermitiana* ou *antihermitiana* se

$$A^H = A \text{ ou } A^H = -A.$$

Claramente, $A = [a_{ij}]$ é hermitiana se, e só se, seus elementos simétricos forem conjugados, ou seja, se cada $a_{ij} = \bar{a}_{ji}$, caso em que cada elemento a_{ii} da diagonal deve ser real. Analogamente, se A for antihermitiana, então cada elemento da diagonal deve ser nulo, $a_{ii} = 0$. (Observe que se $A^H = A$ ou $A^H = -A$, então A necessariamente é quadrada.)

Uma matriz complexa é *unitária* se $AA^H = A^HA = I$, ou seja, se

$$A^H = A^{-1}.$$

Assim, necessariamente, toda matriz unitária é quadrada e invertível. Observamos que uma matriz complexa A é unitária se, e só se, suas linhas (colunas) formam um conjunto ortonormal relativo ao produto escalar de vetores complexos.

Uma matriz complexa A é dita *normal* se comutar com A^H, ou seja, se

$$AA^H = A^HA$$

(Assim, toda matriz normal é quadrada.) Essa definição coincide com a dada para matrizes reais quando A for uma matriz real.

Exemplo 2.16 Consideremos as matrizes complexas seguintes.

$$A = \begin{bmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 5 \end{bmatrix} \qquad B = \frac{1}{2}\begin{bmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 2+3i & 1 \\ i & 1+2i \end{bmatrix}$$

(a) Visualmente, os elementos diagonais de A são reais e os elementos simétricos $1-2i$ e $1+2i$, $4+7i$ e $4-7i$, bem como $-2i$ e $2i$ são, todos, conjugados. Assim, A é hermitiana.
(b) Multiplicar B por B^H fornece I, ou seja, $BB^H = I$. Isso implica que também $B^HB = I$. Assim, $B^H = B^{-1}$, o que significa que B é unitária.
(c) Para mostrar que C é normal, calculamos CC^H e C^HC. Temos

$$CC^H = \begin{bmatrix} 2+3i & 1 \\ i & 1+2i \end{bmatrix}\begin{bmatrix} 2-3i & -i \\ 1 & 1-2i \end{bmatrix} = \begin{bmatrix} 14 & 4-4i \\ 4+4i & 6 \end{bmatrix}$$

e, analogamente, $C^HC = \begin{bmatrix} 14 & 4-4i \\ 4+4i & 6 \end{bmatrix}$. Como $CC^H = C^HC$, essa matriz complexa C é normal.

Observamos que, para matrizes reais, ser hermitiana é o mesmo que ser simétrica, e ser unitária é o mesmo que ser ortogonal.

2.12 MATRIZES EM BLOCOS

Utilizando um sistema de linhas (tracejadas) horizontais e verticais, podemos particionar uma matriz A em submatrizes denominadas *blocos* de A. Certamente podemos dividir uma matriz em blocos de maneiras diferentes. Por exemplo,

$$\left[\begin{array}{cc:cc:c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hdashline 3 & 1 & 4 & 5 & 9 \\ 4 & 6 & -3 & 1 & 8 \end{array}\right], \qquad \left[\begin{array}{cc:ccc} 1 & -2 & 0 & 1 & 3 \\ \hdashline 2 & 3 & 5 & 7 & -2 \\ 3 & 1 & 4 & 5 & 9 \\ \hdashline 4 & 6 & -3 & 1 & 8 \end{array}\right], \qquad \left[\begin{array}{ccc:cc} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hdashline 3 & 1 & 4 & 5 & 9 \\ 4 & 6 & -3 & 1 & 8 \end{array}\right]$$

A conveniência da partição de matrizes em blocos, digamos, A e B, é que o resultado das operações com A e B pode ser obtido fazendo as operações em seus blocos, como se fossem autênticos elementos das matrizes. Ilustramos isso a seguir, onde utilizamos a notação $A = [A_{ij}]$ para uma matriz em blocos A com blocos A_{ij}.

Suponha que $A = [A_{ij}]$ e $B = [B_{ij}]$ sejam matrizes em blocos com o mesmo número de blocos linha e coluna e suponha que blocos correspondentes tenham o mesmo tamanho. Então a soma de blocos correspondentes de A e de B também soma os elementos correspondentes de A e de B e a multiplicação de cada bloco de A pelo mesmo escalar k multiplica cada elemento de A por k. Assim,

$$A + B = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \dots & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & \dots & A_{2n}+B_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \dots & A_{mn}+B_{mn} \end{bmatrix}$$

e

$$kA = \begin{bmatrix} kA_{11} & kA_{12} & \dots & kA_{1n} \\ kA_{21} & kA_{22} & \dots & kA_{2n} \\ \dots & \dots & \dots & \dots \\ kA_{m1} & kA_{m2} & \dots & kA_{mn} \end{bmatrix}$$

O caso da multiplicação matricial é menos óbvio, mas ainda é válido. Isto é, digamos que $U = [U_{ik}]$ e $V = [V_{kj}]$ sejam matrizes em blocos tais que o número de colunas de cada bloco U_{ik} seja igual ao número de linhas de cada bloco V_{kj}, de modo que esteja definido cada produto $U_{ik}V_{kj}$. Então

$$UV = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \dots & \dots & \dots & \dots \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{bmatrix}, \qquad \text{onde} \qquad W_{ij} = U_{i1}V_{1j} + U_{i2}V_{2j} + \cdots + U_{ip}V_{pj}$$

A demonstração dessa fórmula para UV é imediata, mas cheia de detalhes e demorada. Por isso, é deixada como exercício (Problema Complementar 2.86).

Matrizes quadradas em blocos

Seja M uma matriz em blocos. Dizemos que M é uma *matriz quadrada em blocos* se

(i) M é uma matriz quadrada,
(ii) os blocos formam uma matriz quadrada e
(iii) os blocos da diagonal também são matrizes quadradas.

Essas duas últimas condições ocorrem se, e só se, houver o mesmo número de linhas tracejadas horizontais e verticais e se estiverem posicionadas simetricamente.

Considere as duas matrizes em blocos seguintes.

$$A = \left[\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 4 & 4 & 4 & 4 & 4 \\ 3 & 5 & 3 & 5 & 3 \end{array}\right] \qquad \text{e} \qquad B = \left[\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ \hline 3 & 5 & 3 & 5 & 3 \end{array}\right]$$

A matriz em blocos A não é uma matriz quadrada em blocos, porque o segundo e o terceiro blocos da diagonal não são quadrados. No entanto, a matriz em blocos B é uma matriz quadrada em blocos.

Matrizes diagonais em blocos

Seja $M = [A_{ij}]$ uma matriz quadrada em blocos. Dizemos que M é uma *matriz diagonal em blocos* se todos os blocos não diagonais de M forem matrizes nulas, ou seja, $A_{ij} = 0$, com $i \neq j$. Às vezes, denotamos uma tal matriz por

$$M = \text{diag}(A_{11}, A_{22}, ..., A_{rr}) \qquad \text{ou} \qquad M = A_{11} \oplus A_{22} \oplus \cdots \oplus A_{rr}$$

A importância das matrizes diagonais em blocos se deve ao fato de que a álgebra de matrizes em blocos, frequentemente, se reduz à álgebra dos blocos individuais. Especificamente, suponha que $f(x)$ seja um polinômio e que M seja a matriz diagonal em blocos que acabamos de considerar. Então $f(M)$ é uma matriz diagonal em blocos e

$$f(M) = \text{diag}(f(A_{11}), f(A_{22}), ..., f(A_{rr}))$$

Também vale que M é invertível se, e só se, cada A_{ii} é invertível e, nesse caso, M^{-1} é uma matriz diagonal em blocos, com

$$M^{-1} = \text{diag}(A_{11}^{-1}, A_{22}^{-1}, \dots, A_{rr}^{-1})$$

Analogamente, uma matriz quadrada em blocos é dita *matriz triangular superior em blocos* se os blocos abaixo da diagonal forem matrizes nulas e é dita *matriz triangular inferior em blocos* se os blocos acima da diagonal forem matrizes nulas.

Exemplo 2.17 Encontre as matrizes quadradas, triangulares superiores, triangulares ou diagonais dentre as matrizes seguintes.

$$A = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 0 & 6 \end{array}\right], \qquad B = \left[\begin{array}{c|cc|c} 1 & 0 & 0 & 0 \\ \hline 2 & 3 & 4 & 0 \\ 5 & 0 & 6 & 0 \\ \hline 0 & 7 & 8 & 9 \end{array}\right], \qquad C = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 3 \\ 0 & 4 & 5 \end{array}\right], \qquad D = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 6 & 7 \end{array}\right]$$

(a) A é triangular superior, porque o bloco abaixo da diagonal é uma matriz nula.
(b) B é triangular inferior, porque os blocos acima da diagonal são matrizes nulas.
(c) C é diagonal, porque os blocos acima e abaixo da diagonal são matrizes nulas.
(d) D não é nem triangular superior nem triangular inferior. Mais que isso, não há partição de D que a faça ser triangular superior ou inferior em blocos.

Problemas Resolvidos

Soma matricial e multiplicação por escalar

2.1 Dadas $A = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{bmatrix}$, encontre

(a) $A + B$, (b) $2A - 3B$.

(a) Somando os elementos correspondentes, obtemos

$$A + B = \begin{bmatrix} 1+3 & -2+0 & 3+2 \\ 4-7 & 5+1 & -6+8 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 5 \\ -3 & 6 & 2 \end{bmatrix}$$

(b) Primeiro multiplicamos pelos escalares e só depois efetuamos a soma das matrizes.

$$2A - 3B = \begin{bmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{bmatrix} + \begin{bmatrix} -9 & 0 & -6 \\ 21 & -3 & -24 \end{bmatrix} = \begin{bmatrix} -7 & -4 & 0 \\ 29 & 7 & -36 \end{bmatrix}$$

(Observe que multiplicamos B por -3 e depois somamos, em vez de multiplicar B por 3 e depois subtrair. Em geral, isso evita erros.)

2.2 Encontre x, y, z, t tais que $3\begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} x & 6 \\ -1 & 2t \end{bmatrix} + \begin{bmatrix} 4 & x+y \\ z+t & 3 \end{bmatrix}$.

Escrevemos cada lado da igualdade como uma única equação,

$$\begin{bmatrix} 3x & 3y \\ 3z & 3t \end{bmatrix} = \begin{bmatrix} x+4 & x+y+6 \\ z+t-1 & 2t+3 \end{bmatrix}$$

Igualando entradas correspondentes, obtemos o seguinte sistema de quatro equações.

$$3x = x+4, \qquad 3y = x+y+6, \qquad 3z = z+t-1, \qquad 3t = 2t+3$$

ou $\quad 2x = 4, \qquad 2y = 6 + x, \qquad 2z = t - 1, \qquad t = 3$

A solução é $x = 2$, $y = 4$, $z = 1$, $t = 3$.

2.3 Demonstre os itens (i) e (v) do Teorema 2.1. (i) $(A + B) + C = A + (B + C)$, (v) $k(A + B) = kA + kB$.
Suponha que $A = [a_{ij}]$, $B = [b_{ij}]$ e $C = [c_{ij}]$. A demonstração se reduz a mostrar que são iguais as ij-ésimas entradas de cada lado das equações matriciais. [Apenas demonstramos (i) e (v), pois os demais itens do teorema são demonstrados de maneira análoga.]

(i) A ij-ésima entrada de $A + B$ é $a_{ij} + b_{ij}$, portanto, a ij-ésima entrada de $(A + B) + C$ é $(a_{ij} + b_{ij}) + c_{ij}$. Do outro lado, a ij-ésima entrada de $B + C$ é $b_{ij} + c_{ij}$, portanto, a ij-ésima entrada de $A + (B + C)$ é $a_{ij} + (b_{ij} + c_{ij})$. No entanto, para escalares de K, temos

$$(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}).$$

Assim, $(A + B) + C$ e $A + (B + C)$ têm ij-ésimas entradas idênticas, mostrando que $(A + B) + C = A + (B + C)$.

(ii) A ij-ésima entrada de $A + B$ é $a_{ij} + b_{ij}$, portanto, a ij-ésima entrada de $k(A + B)$ é $k(a_{ij} + b_{ij})$. Do outro lado, as ij-ésimas entradas de kA e kB são ka_{ij} e kb_{ij}, respectivamente, portanto, a ij-ésima entrada de $kA + kB$ é $ka_{ij} + kb_{ij}$. No entanto, para escalares de K, temos

$$k(a_{ij} + b_{ij}) = ka_{ij} + kb_{ij}.$$

Assim, $k(A + B)$ e $kA + kB$ têm ij-ésimas entradas idênticas, mostrando que $k(A + B) = kA + kB$.

Multiplicação matricial

2.4 Calcule (a) $[8, -4, 5]\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$, (b) $[6, -1, 7, 5]\begin{bmatrix} 4 \\ -9 \\ -3 \\ 2 \end{bmatrix}$, (c) $[3, 8, -2, 4]\begin{bmatrix} 5 \\ -1 \\ 6 \end{bmatrix}$

(a) Multiplicando as entradas correspondentes e somando, obtemos

$$[8, -4, 5]\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = 8(3) + (-4)(2) + 5(-1) = 24 - 8 - 5 = 11$$

(b) Multiplicando as entradas correspondentes e somando, obtemos

$$[6, -1, 7, 5]\begin{bmatrix} 4 \\ -9 \\ -3 \\ 2 \end{bmatrix} = 24 + 9 - 21 + 10 = 22$$

(c) Esse produto não está definido quando a matriz linha e a matriz coluna têm número de elementos distintos.

2.5 Denotemos uma matriz de tamanho $r \times s$ por $(r \times s)$. Encontre o tamanho dos produtos matriciais que estiverem definidos.

(a) $(2 \times 3)(3 \times 4)$, (c) $(1 \times 2)(3 \times 1)$, (e) $(4 \times 4)(3 \times 3)$

(b) $(4 \times 1)(1 \times 2)$, (d) $(5 \times 2)(2 \times 3)$, (f) $(2 \times 2)(2 \times 4)$

Em cada caso, o produto está definido se os números internos forem iguais e, nesse caso, o tamanho do produto será dado pelos números externos na ordem em que aparecem.

(a) 2×4, (c) Não está definido. (e) Não está definido.

(b) 4×2, (d) 5×3, (f) 2×4

2.6 Sejam $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix}$. Encontre (a) AB, (b) BA.

(a) Como A é uma matriz 2×2 e B é 2×3, o produto AB está definido e é uma matriz 2×3. Para obter as entradas da primeira linha de AB, multiplicamos a primeira linha [1, 3] de A pelas colunas $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\begin{bmatrix} -4 \\ 6 \end{bmatrix}$ de B, respectivamente, como segue.

$$AB = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix} = \begin{bmatrix} 2+9 & 0-6 & -4+18 \\ & & \end{bmatrix} = \begin{bmatrix} 11 & -6 & 14 \\ & & \end{bmatrix}$$

Para obter as entradas da segunda linha de AB, multiplicamos a segunda linha [2, –1] de A pelas colunas de B.

$$AB = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix} = \begin{bmatrix} 11 & -6 & 14 \\ 4-3 & 0+2 & -8-6 \end{bmatrix}$$

Assim,

$$AB = \begin{bmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{bmatrix}.$$

(b) O tamanho de B é 2×3 e o de A é 2×2. Os números internos 3 e 2 não são iguais, portanto, o produto BA não está definido.

2.7 Encontre AB, onde $A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & -2 & 5 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & -1 & 0 & 6 \\ 1 & 3 & -5 & 1 \\ 4 & 1 & -2 & 2 \end{bmatrix}$.

Como A é uma matriz 2×3 e B é 3×4, o produto AB está definido e é uma matriz 2×4. Multiplicamos as linhas de A pelas colunas de B para obter

$$AB = \begin{bmatrix} 4+3-4 & -2+9-1 & 0-15+2 & 12+3-2 \\ 8-2+20 & -4-6+5 & 0+10-10 & 24-2+10 \end{bmatrix} = \begin{bmatrix} 3 & 6 & -13 & 13 \\ 26 & -5 & 0 & 32 \end{bmatrix}.$$

2.8 Encontre (a) $\begin{bmatrix} 1 & 6 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -7 \end{bmatrix}$, (b) $\begin{bmatrix} 2 \\ -7 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ -3 & 5 \end{bmatrix}$, (c) $[2, -7] \begin{bmatrix} 1 & 6 \\ -3 & 5 \end{bmatrix}$.

(a) O primeiro fator é 2×2 e o segundo é 2×1, portanto, o produto está definido como uma matriz 2×1.

$$\begin{bmatrix} 1 & 6 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -7 \end{bmatrix} = \begin{bmatrix} 2-42 \\ -6-35 \end{bmatrix} = \begin{bmatrix} -40 \\ -41 \end{bmatrix}$$

(b) O produto não está definido, porque o primeiro fator é 2×1 e o segundo fator é 2×2.

(c) O primeiro fator é 1×2 e o segundo é 2×2, portanto, o produto está definido como uma matriz linha 1×2.

$$[2, -7] \begin{bmatrix} 1 & 6 \\ -3 & 5 \end{bmatrix} = [2+21, \; 12-35] = [23, -23]$$

2.9 Certamente $0A = 0$ e $A0 = 0$, onde 0 indica matrizes nulas (de tamanhos possivelmente distintos). Encontre matrizes A e B sem entradas nulas tais que $AB = 0$.

Basta tomar $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 6 & 2 \\ -3 & -1 \end{bmatrix}$. Então $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

2.10 Demonstre o item (i) do Teorema 2.2: $(AB)C = A(BC)$.

Sejam $A = [a_{ij}]$, $B = [b_{jk}]$, $C = [c_{kl}]$ e $AB = S = [s_{ik}]$, $BC = T = [t_{jl}]$. Então

$$s_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \qquad \text{e} \qquad t_{jl} = \sum_{k=1}^{n} b_{jk} c_{kl}$$

No produto de $S = AB$ por C, a il-ésima entrada de $(AB)C$ é

$$s_{i1}c_{1l} + s_{i2}c_{2l} + \cdots + s_{in}c_{nl} = \sum_{k=1}^{n} s_{ik}c_{kl} = \sum_{k=1}^{n} \sum_{j=1}^{m} (a_{ij}b_{jk})c_{kl}$$

Por outro lado, no produto de A por $T = BC$, a il-ésima entrada de $A(BC)$ é

$$a_{i1}t_{1l} + a_{i2}t_{2l} + \cdots + a_{in}t_{nl} = \sum_{j=1}^{m} a_{ij}t_{jl} = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{ij}(b_{jk}c_{kl})$$

Essas duas somas são iguais, ou seja, a il-ésima entrada de $(AB)C$ é igual à de $A(BC)$. Assim, $(AB)C = A(BC)$.

2.11 Demonstre o item (ii) do Teorema 2.2: $A(B + C) = AB + AC$.

Sejam $A = [a_{ij}]$, $B = [b_{jk}]$, $C = [c_{jk}]$ e $B + C = D = [d_{ik}]$, $AB = E = [e_{ik}]$, $AC = F = [f_{ik}]$. Então

$$d_{jk} = b_{jk} + c_{jk}, \qquad e_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk}, \qquad f_{ik} = \sum_{j=1}^{m} a_{ij}c_{jk}$$

Assim, a ik-ésima entrada da matriz $AB + AC$ é

$$e_{ik} + f_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk} + \sum_{j=1}^{m} a_{ij}c_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

Por outro lado, a ik-ésima entrada da matriz $AD = A(B + C)$ é

$$a_{i1}d_{1k} + a_{i2}d_{2k} + \cdots + a_{im}d_{mk} = \sum_{j=1}^{m} a_{ij}d_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

Assim, $A(B + C) = AB + AC$, pois as entradas correspondentes são iguais.

Transposta

2.12 Encontre a transposta de cada matriz.

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 7 & 8 & -9 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \qquad C = [1, -3, 5, -7], \qquad D = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$$

Reescrevemos as linhas de cada matriz como colunas para obter as transpostas das matrizes.

$$A^T = \begin{bmatrix} 1 & 7 \\ -2 & 8 \\ 3 & -9 \end{bmatrix}, \qquad B^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \qquad C^T = \begin{bmatrix} 1 \\ -3 \\ 5 \\ -7 \end{bmatrix}, \qquad D^T = [2, -4, 6]$$

(Observe que $B^T = B$ e, portanto, B é dita *simétrica*. Observe, também, que a transposta do vetor linha C é um vetor coluna e que a transposta do vetor coluna D é um vetor linha.)

2.13 Demonstre o item (iv) do Teorema 2.3: $(AB)^T = B^TA^T$.

Sejam $A = [a_{ik}]$ e $B = [b_{kj}]$. Então a ij-ésima entrada de AB é

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}$$

Essa é a ji-ésima entrada (ordem invertida) de $(AB)^T$. Agora a coluna j de B se torna a linha j de B^T e a linha i de A se torna a coluna i de A^T. Logo, a ij-ésima entrada de B^TA^T é

$$[b_{1j}, b_{2j}, \ldots, b_{mj}][a_{i1}, a_{i2}, \ldots, a_{im}]^T = b_{1j}a_{i1} + b_{2j}a_{i2} + \cdots + b_{mj}a_{im}$$

Assim, $(AB)^T = B^TA^T$, pois as entradas correspondentes são iguais.

Matrizes quadradas

2.14 Encontre a diagonal e o traço de cada matriz dada.

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 3 & 6 \\ 2 & -5 & 8 \\ 4 & -2 & 9 \end{bmatrix}, \qquad \text{(b)} \quad B = \begin{bmatrix} 2 & 4 & 8 \\ 3 & -7 & 9 \\ -5 & 0 & 2 \end{bmatrix}, \qquad \text{(c)} \quad C = \begin{bmatrix} 1 & 2 & -3 \\ 4 & -5 & 6 \end{bmatrix}.$$

(a) A diagonal de A consiste nos elementos entre o canto superior esquerdo de A e o canto inferior direito de A ou, em outras palavras, nos elementos a_{11}, a_{22}, a_{33}. Assim, a diagonal de A consiste nos números 1, –5 e 9. O traço de A é a soma dos elementos diagonais. Assim,

$$\operatorname{tr}(A) = 1 - 5 + 9 = 5$$

(b) A diagonal de B consiste nos números 2, –7 e 2. Assim,

$$\operatorname{tr}(B) = 2 - 7 + 2 = -3$$

(c) Só se define diagonal e traço de matrizes quadradas.

2.15 Sejam $A = \begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix}$, $f(x) = 2x^3 - 4x + 5$ e $g(x) = x^2 + 2x + 11$. Encontre

(a) A^2, (b) A^3, (c) $f(A)$, (d) $g(A)$.

(a) $A^2 = AA = \begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix} = \begin{bmatrix} 1+8 & 2-6 \\ 4-12 & 8+9 \end{bmatrix} = \begin{bmatrix} 9 & -4 \\ -8 & 17 \end{bmatrix}$

(b) $A^3 = AA^2 = \begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 9 & -4 \\ -8 & 17 \end{bmatrix} = \begin{bmatrix} 9-16 & -4+34 \\ 36+24 & -16-51 \end{bmatrix} = \begin{bmatrix} -7 & 30 \\ 60 & -67 \end{bmatrix}$

(c) Primeiro substituímos x por A e 5 por $5I$ em $f(x)$, obtendo

$$f(A) = 2A^3 - 4A + 5I = 2\begin{bmatrix} -7 & 30 \\ 60 & -67 \end{bmatrix} - 4\begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix} + 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Em seguida, multiplicamos pelos escalares e depois efetuamos a soma matricial.

$$f(A) = \begin{bmatrix} -14 & 60 \\ 120 & -134 \end{bmatrix} + \begin{bmatrix} -4 & -8 \\ -16 & 12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -13 & 52 \\ 104 & -117 \end{bmatrix}$$

(d) Substituindo x por A e 11 por $11I$ em $g(x)$, calculamos

$$g(A) = A^2 + 2A - 11I = \begin{bmatrix} 9 & -4 \\ -8 & 17 \end{bmatrix} + 2\begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix} - 11\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 9 & -4 \\ -8 & 17 \end{bmatrix} + \begin{bmatrix} 2 & 4 \\ 8 & -6 \end{bmatrix} + \begin{bmatrix} -11 & 0 \\ 0 & -11 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Como $g(A)$ é a matriz nula, A é uma raiz do polinômio $g(x)$.

2.16 Seja $A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$. (a) Encontre um vetor coluna não nulo $u = \begin{bmatrix} x \\ y \end{bmatrix}$ tal que $Au = 3u$. (b) Descreva todos esses vetores coluna.

(a) Primeiro montamos a equação matricial $Au = 3u$ e depois escrevemos cada lado como uma única matriz (vetor coluna), como segue.

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 3\begin{bmatrix} x \\ y \end{bmatrix}, \qquad \text{e, então} \qquad \begin{bmatrix} x+3y \\ 4x-3y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix}$$

Igualamos as entradas correspondentes para obter um sistema de equações.

$$\begin{aligned} x + 3y &= 3x \\ 4x - 3y &= 3y \end{aligned} \qquad \text{ou} \qquad \begin{aligned} 2x - 3y &= 0 \\ 4x - 6y &= 0 \end{aligned} \qquad \text{ou} \qquad 2x - 3y = 0$$

Esse sistema reduziu a uma equação linear não degenerada com duas incógnitas, portanto, tem uma infinidade de soluções. Para obter uma solução não nula, tomamos, por exemplo, $y = 2$; então $x = 3$. Assim, $u = (3, 2)^T$ é um vetor não nulo tal que $Au = 3u$.

(b) Para obter a solução geral, tomamos $y = a$, onde a é um parâmetro. Substituindo $y = a$ em $2x - 3y = 0$, obtemos $x = \frac{3}{2}a$. Assim, $u = (\frac{3}{2}a, a)^T$ representa todas as soluções.

Matrizes invertíveis, inversas

2.17 Mostre que $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$ e $B = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$ são inversas.

Para isso, calculamos o produto AB

$$AB = \begin{bmatrix} -11+0+12 & 2+0-2 & 2+0-2 \\ -22+4+18 & 4+0-3 & 4-1-3 \\ -44-4+48 & 8+0-8 & 8+1-8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

Como $AB = I$, podemos concluir (Teorema 3.18) que $BA = I$. Por isso, A e B são inversas.

2.18 Encontre, caso exista, a inversa de cada matriz dada.

(a) $A = \begin{bmatrix} 5 & 3 \\ 4 & 2 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & -3 \\ 1 & 3 \end{bmatrix}$, (c) $\begin{bmatrix} -2 & 6 \\ 3 & -9 \end{bmatrix}$.

Utilize a fórmula da inversa de matrizes de ordem 2 vista na Seção 2.9.

(a) Primeiro calculamos $|A| = 5(2) - 3(4) = 10 - 12 = -2$. Em seguida, trocamos de lugar os dois elementos da diagonal, trocamos o sinal dos elementos fora da diagonal e multiplicamos por $1/|A|$, como segue.

$$A^{-1} = -\frac{1}{2}\begin{bmatrix} 2 & -3 \\ -4 & 5 \end{bmatrix} = \begin{bmatrix} -1 & \frac{3}{2} \\ 2 & -\frac{5}{2} \end{bmatrix}$$

(b) Primeiro calculamos $|B| = 2(3) - (-3)(1) = 6 + 3 = 9$. Em seguida, trocamos de lugar os dois elementos da diagonal, trocamos o sinal dos elementos fora da diagonal e multiplicamos por $1/|B|$, como segue.

$$B^{-1} = \frac{1}{9}\begin{bmatrix} 3 & 3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

(c) Primeiro calculamos $|C| = -2(-9) - 6(3) = 18 - 18 = 0$. Como $|C| = 0$, essa matriz não tem inversa.

2.19 Seja $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{bmatrix}$. Encontre $A^{-1} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$.

Multiplicando A por A^{-1} e igualando as nove entradas do produto com as nove entradas da matriz identidade I, obtemos o sistema de três equações com três incógnitas seguinte.

$$\begin{array}{lll} x_1 + y_1 + z_1 = 1 & x_2 + y_2 + z_2 = 0 & x_3 + y_3 + z_3 = 0 \\ y_1 + 2z_1 = 0 & y_2 + 2z_2 = 1 & y_3 + 2z_3 = 0 \\ x_1 + 2y_1 + 4z_1 = 0 & x_2 + 2y_2 + 4z_2 = 0 & x_3 + 2y_3 + 4z_3 = 1 \end{array}$$

[Observe que a matriz de coeficientes dos três sistemas é A.]
Resolvendo os três sistemas nas nove incógnitas, obtermos

$$x_1 = 0, \quad y_1 = 2, \quad z_1 = -1; \qquad x_2 = -2, \quad y_2 = 3, \quad z_2 = -1; \qquad x_3 = 1, \quad y_3 = -2, \quad z_3 = 1$$

Assim,
$$A^{-1} = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix}$$

(**Observação:** No Capítulo 3 veremos uma maneira eficiente de resolver os três sistemas.)

2.20 Sejam A e B matrizes invertíveis (de mesmo tamanho). Mostre que AB também é invertível e que $(AB)^{-1} = B^{-1}A^{-1}$. [Assim, por indução, $(A_1A_2 \ldots A_m)^{-1} = A_m^{-1} \ldots A_2^{-1}A_1^{-1}$.]

Usando a associatividade do produto matricial, obtemos

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = A^{-1}IB = B^{-1}B = I$$

Assim, $(AB)^{-1} = B^{-1}A^{-1}$.

Matrizes diagonais e triangulares

2.21 Escreva em forma matricial as matrizes $A = \text{diag}(4, -3, 7)$, $B = \text{diag}(2, -6)$ e $C = \text{diag}(3, -8, 0, 5)$.

Colocamos os escalares dados na diagonal e completamos com 0.

$$\text{A} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \qquad \text{B} = \begin{bmatrix} 2 & 0 \\ 0 & -6 \end{bmatrix}, \qquad \text{C} = \begin{bmatrix} 3 & & & \\ & -8 & & \\ & & 0 & \\ & & & 5 \end{bmatrix}$$

2.22 Sejam $A = \text{diag}(2, 3, 5)$ e $B = \text{diag}(7, 0, -4)$. Encontre

(a) AB, A^2, B^2; (b) $f(A)$, com $f(x) = x^2 + 3x - 2$; (c) A^{-1} e B^{-1}.

(a) A matriz produto AB é uma matriz diagonal obtida multiplicando as entradas diagonais correspondentes, ou seja,

$$AB = \text{diag}(2(7), 3(0), 5(-4)) = \text{diag}(14, 0, -20).$$

Os quadrados A^2 e B^2 são obtidos tomando o quadrado de cada entrada diagonal, ou seja,

$$A^2 = \text{diag}(2^2, 3^2, 5^2) = \text{diag}(4, 9, 25) \qquad \text{e} \qquad B^2 = \text{diag}(49, 0, 16)$$

(b) $f(A)$ é uma matriz diagonal obtida calculando $f(x)$ em cada entrada diagonal de A. Temos

$$f(2) = 4 + 6 - 2 = 8, \qquad f(3) = 9 + 9 - 2 = 16, \qquad f(5) = 25 + 15 - 2 = 38$$

Assim, $f(A) = \text{diag}(8, 16, 38)$.

(c) A inversa de uma matriz diagonal é uma matriz diagonal tomando o recíproco de cada elemento diagonal. Assim, $A^{-1} = \text{diag}(\frac{1}{2}, \frac{1}{3}, \frac{1}{5})$, mas B não tem inversa por possuir um 0 na diagonal.

2.23 Encontre uma matriz A de tamanho 2×2 tal que A^2 seja diagonal, mas não A.

Tomando $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, temos $A^2 = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$, que é diagonal.

2.24 Encontre uma matriz triangular superior A tal que $A^3 = \begin{bmatrix} 8 & -57 \\ 0 & 27 \end{bmatrix}$.

Seja $A = \begin{bmatrix} x & y \\ 0 & z \end{bmatrix}$. Então $x^3 = 8$, portanto, $x = 2$ e $z^3 = 27$, portanto, $z = 3$. Em seguida, calculamos A^3 usando $x = 2$ e $z = 3$, obtendo

$$A^2 = \begin{bmatrix} 2 & y \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2 & y \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 5y \\ 0 & 9 \end{bmatrix} \qquad \text{e} \qquad A^3 = \begin{bmatrix} 2 & y \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 4 & 5y \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 8 & 19y \\ 0 & 27 \end{bmatrix}$$

Assim, $19y = -57$, ou $y = -3$. Isso mostra que $A = \begin{bmatrix} 2 & -3 \\ 0 & 3 \end{bmatrix}$.

2.25 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes triangulares superiores. Demonstre que AB é triangular superior com diagonal $a_{11}b_{11}, a_{22}b_{22}, \ldots, a_{nn}b_{nn}$.

Seja $AB = [c_{ij}]$. Então $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ e $c_{ii} = \sum_{k=1}^{n} a_{ik}b_{ki}$. Suponha que $i > j$. Então, dado qualquer k, ou $i > k$, ou $k > j$, portanto, ou $a_{ik} = 0$, ou $b_{kj} = 0$. Segue que $c_{ij} = 0$ e AB é triangular superior. Suponha que $i = j$. Então, dado qualquer k, de $k < i$ decorre $a_{ik} = 0$ e, de $k > i$, decorre $b_{ki} = 0$. Segue que $c_{ii} = a_{ii}b_{ii}$, conforme afirmado. [Isso demonstra uma parte do item (i) do Teorema 2.5; as afirmações para $A + B$ e kA são deixadas como exercícios.]

Matrizes reais especiais: simétricas e ortogonais

2.26 Decida se as matrizes dadas são simétricas (ou seja, iguais à sua transposta) ou antissimétricas (ou seja, iguais ao simétrico de sua transposta).

(a) $A = \begin{bmatrix} 5 & -7 & 1 \\ -7 & 8 & 2 \\ 1 & 2 & -4 \end{bmatrix}$, (b) $B = \begin{bmatrix} 0 & 4 & -3 \\ -4 & 0 & 5 \\ 3 & -5 & 0 \end{bmatrix}$, (c) $C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(a) Visualmente, os elementos simétricos (ou seja, imagens espelhadas pela diagonal) de A são –7 e –7, 1 e 1, 2 e 2. Assim, A é simétrica pela igualdade de seus elementos simétricos.

(b) Visualmente, os elementos da diagonal são todos nulos e os elementos simétricos de B, a saber, 4 e –4, –3 e 3, 5 e –5, são dois a dois simétricos. Assim, B é antissimétrica.

(c) Como C não é quadrada, C não é nem simétrica nem antissimétrica.

2.27 Suponha que $B = \begin{bmatrix} 4 & x+2 \\ 2x-3 & x+1 \end{bmatrix}$ seja simétrica. Encontre x e B.

Igualando os elementos simétricos $x + 2$ e $2x - 3$, obtemos $2x - 3 = x + 2$, ou $x = 5$. Logo, $B = \begin{bmatrix} 4 & 7 \\ 7 & 6 \end{bmatrix}$.

2.28 Seja A uma matriz 2×2 real arbitrária.

(a) Demonstre que se (a, b) for a primeira linha de A, então $a^2 + b^2 = 1$ e

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}.$$

(b) Demonstre o Teorema 2.7. Para algum número real θ, temos

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

(a) Suponha que (x, y) seja a segunda linha de A. Como as linhas de A constituem um conjunto ortonormal, temos

$$a^2 + b^2 = 1, \qquad x^2 + y^2 = 1, \qquad ax + by = 0$$

Analogamente, as colunas constituem um conjunto ortonormal, portanto

$$a^2 + x^2 = 1, \qquad b^2 + y^2 = 1, \qquad ab + xy = 0$$

Segue que $x^2 = 1 - a^2 = b^2$, de modo que $x = \pm b$.

Caso (i): $x = b$. Então $b(a + y) = 0$, logo $y = -a$.

Caso (ii): $x = -b$. Então $b(y - a) = 0$, logo $y = a$.

Disso decorre que, conforme afirmado,

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

(b) Como $a^2 + b^2 = 1$, temos $-1 \leq a \leq 1$. Seja $a = \cos\theta$. Então $b^2 = 1 - \cos^2\theta$, de modo que $b = \operatorname{sen}\theta$. Assim demonstramos o teorema.

2.29 Encontre uma matriz ortogonal A de tamanho 2×2 cuja primeira linha seja um múltiplo (positivo) de (3, 4).

Normalizando (3, 4), obtemos $(\frac{3}{5}, \frac{4}{5})$. Então, pelo Problema 2.28,

$$A = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}.$$

2.30 Encontre uma matriz 3×3 ortogonal P cujas duas primeiras linhas sejam múltiplos de $u_1 = (1, 1, 1)$ e $u_2 = (0, -1, 1)$, respectivamente. (Observe que u_1 e u_2 são ortogonais, conforme exige a definição de matriz ortogonal.)

Inicialmente procuramos encontrar algum vetor não nulo u_3 que seja ortogonal a u_1 e u_2, por exemplo, seu produto vetorial $u_3 = u_1 \times u_2 = (2, -1, -1)$. Sejam A matriz de linhas u_1, u_2 e u_3 e P a matriz obtida de A normalizando cada linha. Assim,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 2 & -1 & -1 \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}$$

Matrizes complexas: hermitianas e unitárias

2.31 Encontre A^H com (a) $A = \begin{bmatrix} 3-5i & 2+4i \\ 6+7i & 1+8i \end{bmatrix}$, (b) $A = \begin{bmatrix} 2-3i & 5+8i \\ -4 & 3-7i \\ -6-i & 5i \end{bmatrix}$

Por definição, $A^H = \bar{A}^T$ é a transposta conjugada de A. Assim,

(a) $A^H = \begin{bmatrix} 3+5i & 6-7i \\ 2-4i & 1-8i \end{bmatrix}$, (b) $A^H = \begin{bmatrix} 2+3i & -4 & -6+i \\ 5-8i & 3+7i & -5i \end{bmatrix}$

2.32 Mostre que $A = \begin{bmatrix} \frac{1}{3} - \frac{2}{3}i & \frac{2}{3}i \\ -\frac{2}{3}i & -\frac{1}{3} - \frac{2}{3}i \end{bmatrix}$ é unitária.

As linhas de A constituem um conjunto ortonormal, como segue.

$$\left(\frac{1}{3} - \frac{2}{3}i, \frac{2}{3}i\right) \cdot \left(\frac{1}{3} - \frac{2}{3}i, \frac{2}{3}i\right) = \left(\frac{1}{9} + \frac{4}{9}\right) + \frac{4}{9} = 1$$

$$\left(\frac{1}{3} - \frac{2}{3}i, \frac{2}{3}i\right) \cdot \left(-\frac{2}{3}i, -\frac{1}{3} - \frac{2}{3}i\right) = \left(\frac{2}{9}i + \frac{4}{9}\right) + \left(-\frac{2}{9}i - \frac{4}{9}\right) = 0$$

$$\left(-\frac{2}{3}i, -\frac{1}{3} - \frac{2}{3}i\right) \cdot \left(-\frac{2}{3}i, -\frac{1}{3} - \frac{2}{3}i\right) = \frac{4}{9} + \left(\frac{1}{9} + \frac{4}{9}\right) = 1$$

Assim, A é unitária.

2.33 Demonstre o análogo complexo do Teorema 2.6. Seja A uma matriz complexa. As afirmações seguintes são equivalentes. (i) A é unitária. (ii) As linhas de A constituem um conjunto ortonormal. (iii) As colunas de A constituem um conjunto ortonormal.

(A demonstração é quase idêntica à fornecida no Exemplo 2.13 para o caso de uma matriz real A de tamanho 3×3.) Recordemos que os vetores $u_1, u_2, ..., u_n$ de $\mathbf{C}^n$ constituem um conjunto ortonormal se forem vetores unitários e dois a dois ortogonais segundo o produto escalar de $\mathbf{C}^n$ definido por

$$(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) = a_1\bar{b}_1 + a_2\bar{b}_2 + \cdots + a_n\bar{b}_n$$

Seja A uma matriz unitária com linhas $R_1, R_2, ..., R_n$. Então $\bar{R}_1^T, \bar{R}_2^T, \ldots, \bar{R}_n^T$ são as colunas de A^H. Denotemos $AA^H = [c_{ij}]$. Por multiplicação matricial, $c_{ij} = R_i\bar{R}_j^T = R_i \cdot R_j$. Como A é unitária, temos $AA^H = I$. Multiplicando A por A^H e igualando cada entrada c_{ij} com a correspondente entrada de I, obtemos as n^2 equações seguintes.

$$R_1 \cdot R_1 = 1, \quad R_2 \cdot R_2 = 1, \quad \ldots, \quad R_n \cdot R_n = 1, \qquad \text{e} \qquad R_i \cdot R_j = 0, \text{ com } i \neq j$$

Portanto, as linhas de A são vetores unitários e dois a dois ortogonais, de modo que formam um conjunto ortonormal de vetores. Analogamente, a condição $A^TA = I$ implica que as colunas de A formam um conjunto ortonormal de vetores. Além disso, como podemos retornar a cada passo, também valem as recíprocas, demonstrando o teorema.

Matrizes em blocos

2.34 Considere as matrizes em blocos (que são partições da mesma matriz) a seguir.

$$\text{(a)} \left[\begin{array}{cc|c|c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right], \qquad \text{(b)} \left[\begin{array}{cc|ccc} 1 & -2 & 0 & 1 & 3 \\ \hline 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right]$$

Encontre o tamanho de cada matriz em blocos e também o de cada bloco matricial.

(a) A matriz em blocos tem duas linhas de matrizes e três colunas de matrizes, portanto, é de tamanho 2×3. Os blocos têm tamanhos 2×2, 2×2 e 2×1 na primeira linha e 1×2, 1×2 e 1×1 na segunda linha.

(b) A matriz em blocos tem tamanho 3×2 e os blocos têm tamanhos 1×3 e 1×2 em cada uma das três linhas.

2.35 Calcule AB usando multiplicação em blocos, onde

$$A = \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right] \qquad \text{e} \qquad B = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Aqui, $A = \begin{bmatrix} E & F \\ 0_{1\times 2} & G \end{bmatrix}$ e $B = \begin{bmatrix} R & S \\ 0_{1\times 3} & T \end{bmatrix}$, onde E, F, G, R, S e T são os blocos dados e $0_{1\times 2}$ e $0_{1\times 3}$ são matrizes nulas com os tamanhos indicados. Assim,

$$AB = \begin{bmatrix} ER & ES + FT \\ 0_{1\times 3} & GT \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{bmatrix} & \begin{bmatrix} 3 \\ 7 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ [0 \quad 0 \quad 0] & 2 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 15 & 4 \\ 19 & 26 & 33 & 7 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

2.36 Seja $M = \text{diag}(A, B, C)$, com $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = [5]$, $C = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}$. Encontre M^2.

Como M é diagonal em blocos, elevamos cada bloco ao quadrado

$$A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}, \qquad B^2 = [25], \qquad C^2 = \begin{bmatrix} 16 & 24 \\ 40 & 64 \end{bmatrix},$$

e obtemos

$$M^2 = \left[\begin{array}{cc|c|cc} 7 & 10 & & & \\ 15 & 22 & & & \\ \hline & & 25 & & \\ \hline & & & 16 & 24 \\ & & & 40 & 64 \end{array}\right]$$

Polinômios matriciais

2.37 Sejam $f(x)$ e $g(x)$ polinômios e A uma matriz quadrada. Mostre que

(a) $(f + g)(A) = f(A) + g(A)$,

(b) $(f \cdot g)(A) = f(A)g(A)$,

(c) $f(A)g(A) = g(A)f(A)$.

Suponha que $f(x) = \sum_{i=1}^{r} a_i x^i$ e $g(x) = \sum_{j=1}^{s} b_j x^j$.

(a) Adicionando potências de x com coeficientes 0, podemos supor que $r = s = n$. Então

$$f(x) + g(x) = \sum_{i=1}^{n} (a_i + b_i) x^i$$

Logo, $$(f + g)(A) = \sum_{i=1}^{n} (a_i + b_i) A^i = \sum_{i=1}^{n} a_i A^i + \sum_{i=1}^{n} b_i A^i = f(A) + g(A)$$

(b) Temos $f(x)g(x) = \sum_{i,j} a_i b_j x^{i+j}$. Então

$$f(A)g(A) = \left(\sum_i a_i A^i \right) \left(\sum_j b_j A^j \right) = \sum_{i,j} a_i b_j A^{i+j} = (fg)(A)$$

(c) Usando $f(x)g(x) = g(x)f(x)$, vemos que

$$f(A)g(A) = (fg)(A) = (gf)(A) = g(A)f(A)$$

Problemas Complementares

Álgebra matricial

Os Problemas 2.38 a 2.41 se referem às matrizes seguintes.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 0 \\ -6 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -3 & 4 \\ 2 & 6 & -5 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 7 & -1 \\ 4 & -8 & 9 \end{bmatrix}$$

2.38 Encontre (a) $5A - 2B$, (b) $2A + 3B$, (c) $2C - 3D$.

2.39 Encontre (a) AB e $(AB)C$, (b) BC e $A(BC)$. [Lembre que $(AB)C = A(BC)$.]

2.40 Encontre (a) A^2 e A^3, (b) AD e BD, (c) CD.

2.41 Encontre (a) A^T, (b) B^T, (c) $(AB)^T$, (d) A^TB^T. [Lembre que $A^TB^T \neq (AB)^T$.]

Os Problemas 2.42 e 2.43 se referem às matrizes seguintes.

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 0 & -3 \\ -1 & -2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -3 & 0 & 1 \\ 5 & -1 & -4 & 2 \\ -1 & 0 & 0 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.$$

2.42 Encontre (a) $3A - 4B$, (b) AC, (c) BC, (d) AD, (e) BD, (f) CD.

2.43 Encontre (a) A^T, (b) A^TB, (c) A^TC.

2.44 Seja $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. Encontre uma matriz B de tamanho 2×3 com entradas não nulas e tal que $AB = 0$.

2.45 Sejam $e_1 = [1, 0, 0]$, $e_2 = [0, 1, 0]$, $e_3 = [0, 0, 1]$ e $A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix}$. Encontre e_1A, e_2A, e_3A.

2.46 Seja $e_i = [0, ..., 0, 1, 0, ..., 0]$, em que o 1 é a i-ésima entrada. Mostre que

(a) $e_iA = A_i$, a i-ésima linha de A.

(b) $Be_j^T = B^j$, a j-ésima coluna de B.

(c) Se $e_iA = e_iB$ para cada i, então $A = B$.

(d) Se $Ae_j^T = Be_j^T$ para cada j, então $A = B$.

2.47 Demonstre os itens (iii) e (iv) do Teorema 2.2: (iii) $(A + B)C = AC + BC$, (iv) $k(AB) = (kA)B = A(kB)$.

2.48 Demonstre o Teorema 2.3: (i) $(A + B)^T = A^T + B^T$, (ii) $(A^T)^T = A$, (iii) $(kA)^T = kA^T$.

2.49 Demonstre as afirmações seguintes.

(a) Se A tem uma linha nula, então AB tem uma linha nula.

(b) Se B tem uma coluna nula, então AB tem uma coluna nula.

Matrizes quadradas, inversas

2.50 Encontre a diagonal e o traço de cada uma das matrizes seguintes.

(a) $A = \begin{bmatrix} 2 & -5 & 8 \\ 3 & -6 & -7 \\ 4 & 0 & -1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 3 & -4 \\ 6 & 1 & 7 \\ 2 & -5 & -1 \end{bmatrix}$, (c) $C = \begin{bmatrix} 4 & 3 & -6 \\ 2 & -5 & 0 \end{bmatrix}$

Os Problemas 2.51 a 2.53 se referem às matrizes $A = \begin{bmatrix} 2 & -5 \\ 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -2 \\ 1 & -6 \end{bmatrix}$, $C = \begin{bmatrix} 6 & -4 \\ 3 & -2 \end{bmatrix}$.

2.51 Encontre (a) A^2 e A^3, (b) $f(A)$ e $g(A)$, onde

$$f(x) = x^3 - 2x^2 - 5, \qquad g(x) = x^2 - 3x + 17.$$

2.52 Encontre (a) B^2 e B^3, (b) $f(B)$ e $g(B)$, onde

$$f(x) = x^2 + 2x - 22, \; g(x) = x^2 - 3x - 6.$$

2.53 Encontre um vetor coluna não nulo v tal que $Cv = 4v$.

2.54 Encontre a inversa de cada uma das matrizes seguintes (se existir).

$$A = \begin{bmatrix} 7 & 4 \\ 5 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, \qquad C = \begin{bmatrix} 4 & -6 \\ -2 & 3 \end{bmatrix}, \qquad D = \begin{bmatrix} 5 & -2 \\ 6 & -3 \end{bmatrix}$$

2.55 Encontre as inversas de $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 1 & 3 & 7 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 3 & -2 \end{bmatrix}$. [*Sugestão*: ver Problema 2.19.]

2.56 Suponha que A seja invertível. Mostre que, se $AB = AC$, então $B = C$. Dê um exemplo de uma matriz A não nula tal que $AB = AC$, mas $B \neq C$.

2.57 Encontre matrizes 2×2 invertíveis A e B tais que $A + B \neq 0$ e $A + B$ não seja invertível.

2.58 Demonstre as afirmações seguintes.

(a) A é invertível se, e só se, A^T é invertível.

(b) As operações de inversão e transposição comutam, ou seja, $(A^T)^{-1} = (A^{-1})^T$.

(c) Se A tem alguma linha nula ou coluna nula, então A não é invertível.

Matrizes diagonais e triangulares

2.59 Sejam $A = \text{diag}(1, 2, -3)$ e $B = \text{diag}(2, -5, 0)$. Encontre (a) AB, A^2 e B^2; (b) $f(A)$, onde $f(x) = x^2 + 4x - 3$; (c) A^{-1} e B^{-1}.

2.60 Sejam $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. (a) Encontre A^n. (b) Encontre B^n.

2.61 Encontre todas as matrizes triangulares reais A tais que $A^2 = B$, com (a) $B = \begin{bmatrix} 4 & 21 \\ 0 & 25 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 4 \\ 0 & -9 \end{bmatrix}$.

2.62 Seja $A = \begin{bmatrix} 5 & 2 \\ 0 & k \end{bmatrix}$. Encontre todos números k tais que A seja uma raiz do polinômio

(a) $f(x) = x^2 - 7x + 10$; (b) $g(x) = x^2 - 25$; (c) $h(x) = x^2 - 4$.

2.63 Seja $B = \begin{bmatrix} 1 & 0 \\ 26 & 27 \end{bmatrix}$. Encontre uma matriz A tal que $A^3 = B$.

2.64 Seja $B = \begin{bmatrix} 1 & 8 & 5 \\ 0 & 9 & 5 \\ 0 & 0 & 4 \end{bmatrix}$. Encontre uma matriz triangular A com entradas diagonais positivas e tal que $A^2 = B$.

2.65 Usando somente os elementos 0 e 1, encontre o número de matrizes 3×3 que sejam (a) diagonais, (b) triangulares superiores, (c) invertíveis e triangulares superiores. Generalize para matrizes $n \times n$.

2.66 Seja $D_k = kI$ a matriz escalar correspondente ao escalar k. Mostre que

(a) $D_k A = kA$, (b) $BD_k = kB$, (c) $D_k + D_{k'} = D_{k+k'}$, (d) $D_k D_{k'} = D_{kk'}$

2.67 Suponha que $AB = C$, onde A e C são matrizes triangulares superiores.

(a) Encontre matrizes 2×2 não nulas A, B, C, em que B não seja triangular superior.

(b) Suponha que A também seja invertível. Mostre que B também deve ser triangular superior.

Tipos especiais de matrizes reais

2.68 Encontre x, y, z tais que A seja simétrica, sendo

(a) $A = \begin{bmatrix} 2 & x & 3 \\ 4 & 5 & y \\ z & 1 & 7 \end{bmatrix}$, (b) $A = \begin{bmatrix} 7 & -6 & 2x \\ y & z & -2 \\ x & -2 & 5 \end{bmatrix}$.

2.69 Seja A uma matriz quadrada. Mostre que (a) $A + A^T$ é simétrica, (b) $A - A^T$ é antissimétrica, (c) $A = B + C$, em que B é simétrica e C é antissimétrica.

2.70 Escreva $A = \begin{bmatrix} 4 & 5 \\ 1 & 3 \end{bmatrix}$ como a soma de uma matriz simétrica B e uma matriz antissimétrica C.

2.71 Suponha que A e B sejam simétricas. Mostre que as matrizes seguintes também são simétricas.

(a) $A + B$; (b) kA, com qualquer escalar k; (c) A^2;

(d) A^n, com $n > 0$; (e) $f(A)$, com qualquer polinômio $f(x)$.

2.72 Encontre uma matriz 2×2 ortogonal P cuja primeira linha seja algum múltiplo de

(a) (3, –4), (b) (1, 2)

2.73 Encontre uma matriz 3×3 ortogonal P cujas duas primeiras linhas sejam múltiplos de

(a) (1, 2, 3) e (0, –2, 3), (b) (1, 3, 1) e (1, 0, –1).

2.74 Suponha que A e B sejam ortogonais. Mostre que A^T, A^{-1} e AB também são ortogonais.

2.75 Decida quais das matrizes seguintes são normais. $A = \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Matrizes complexas

2.76 Encontre números reais x, y, z tais que A seja hermitiana, com $A = \begin{bmatrix} 3 & x + 2i & yi \\ 3 - 2i & 0 & 1 + zi \\ yi & 1 - xi & -1 \end{bmatrix}$.

2.77 Suponha que A seja uma matriz complexa. Mostre que AA^H e A^HA são hermitianas.

2.78 Seja A uma matriz quadrada. Mostre que (a) $A + A^H$ é hermitiana, (b) $A - A^H$ é antihermitiana, (c) $A = B + C$, em que B é hermitiana e C é antihermitiana.

2.79 Decida quais das matrizes seguintes são unitárias.

$$A = \begin{bmatrix} i/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -i/2 \end{bmatrix}, \qquad B = \frac{1}{2}\begin{bmatrix} 1 + i & 1 - i \\ 1 - i & 1 + i \end{bmatrix}, \qquad C = \frac{1}{2}\begin{bmatrix} 1 & -i & -1 + i \\ i & 1 & 1 + i \\ 1 + i & -1 + i & 0 \end{bmatrix}$$

2.80 Suponha que A e B sejam unitárias. Mostre que A^H, A^{-1} e AB são unitárias.

2.81 Decida quais das matrizes seguintes são normais. $A = \begin{bmatrix} 3 + 4i & 1 \\ i & 2 + 3i \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 0 \\ 1 - i & i \end{bmatrix}$.

Matrizes em blocos

2.82 Sejam $U = \left[\begin{array}{cc|ccc} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ \hline 0 & 0 & 5 & 1 & 2 \\ 0 & 0 & 3 & 4 & 1 \end{array}\right]$ e $V = \left[\begin{array}{cc|cc} 3 & -2 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & -4 & 1 \end{array}\right]$.

(a) Encontre UV usando multiplicação em blocos. (b) Serão U e V matrizes diagonais em blocos?

(c) Será UV diagonal em blocos?

2.83 Particione cada uma das seguintes matrizes de tal modo que seja uma matriz quadrada em blocos com o maior número possível de blocos na diagonal.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

2.84 Encontre M^2 e M^3 no caso (a) $M = \left[\begin{array}{c|cc|c} 2 & 0 & 0 & 0 \\ \hline 0 & 1 & 4 & 0 \\ 0 & 2 & 1 & 0 \\ \hline 0 & 0 & 0 & 3 \end{array}\right]$, (b) $M = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & 4 & 5 \end{array}\right]$.

2.85 Para cada matriz do Problema 2.84, encontre $f(M)$, sendo $f(x) = x^2 + 4x - 5$.

2.86 Sejam $U = [U_{ik}]$ e $V = [V_{kj}]$ matrizes em blocos para as quais esteja definido o produto UV e tais que o número de colunas de cada bloco U_{ik} seja igual ao número de linhas de cada bloco V_{kj}. Mostre que $UV = [W_{ij}]$, com $W_{ij} = \sum_k U_{ik}V_{kj}$.

2.87 Suponha que M e N sejam matrizes diagonais em blocos tais que blocos correspondentes tenham o mesmo tamanho, digamos, $M = \text{diag}(A_i)$ e $N = \text{diag}(B_i)$. Mostre que

(i) $M + N = \text{diag}(A_i + B_i)$, (iii) $MN = \text{diag}(A_iB_i)$,

(ii) $kM = \text{diag}(kA_i)$, (iv) $f(M) = \text{diag}(f(A_i))$, para qualquer polinômio $f(x)$.

Respostas dos Problemas Complementares

Notação $A = [R_1; R_2; \ldots]$ denota uma matriz A com linhas $R_1, R_2, \ldots$

2.38 (a) $[-5, 10;\quad 27, -34]$, (b) $[17, 4;\quad -12, 13]$, (c) $[-7, -27, 11;\quad -8, 36, -37]$

2.39 (a) $[-7, 14;\quad 39, -28]$, $[21, 105, -98;\quad -17, -285, 296]$
(b) $[5, -15, 20;\quad 8, 60, -59]$, $[21, 105, -98;\quad -17, -285, 296]$

2.40 (a) $[7, -6;\quad -9, 22]$, $[-11, 38;\quad 57, -106]$;
(b) $[11, -9, 17;\quad -7, 53, -39]$, $[15, 35, -5;\quad 10, -98, 69]$; (c) Não está definido.

2.41 (a) $[1, 3;\quad 2, -4]$, (b) $[5, -6;\quad 0, 7]$, (c) $[-7, 39;\quad 14, -28]$, (d) $[5, 15;\quad 10, -40]$

2.42 (a) $[-13, -3, 18;\quad 4, 17, 0]$, (b) $[-5, -2, 4, 5;\quad 11, -3, -12, 18]$,
(c) $[11, -12, 0, -5;\quad -15, 5, 8, 4]$, (d) $[9;\quad 9]$, (e) $[-1;\quad 9]$, (f) Não está definido.

2.43 (a) $[1, 0; \quad -1, 3; \quad 2, 4]$, (b) $[4, 0, -3; \quad -7, -6, 12; \quad 4, -8, 6]$, (c) Não está definido.

2.44 $[2, 4, 6; \quad -1, -2, -3]$

2.45 $[a_1, a_2, a_3, a_4]$, $[b_1, b_2, b_3, b_4]$, $[c_1, c_2, c_3, c_4]$

2.50 (a) $2, -6, -1, \mathrm{tr}(A) = -5$, (b) $1, 1, -1, \mathrm{tr}(B) = 1$, (c) Não estão definidos.

2.51 (a) $[-11, -15; \quad 9, -14]$, $[-67, 40; \quad -24, -59]$, (b) $[-50, 70; \quad -42, -36]$, $g(A) = 0$

2.52 (a) $[14, 4; \quad -2, 34]$, $[60, -52; \quad 26, -200]$, (b) $f(B) = 0$, $[-4, 10; \quad -5, 46]$

2.53 $u = [2a, a]^T$

2.54 $[3, -4; \quad -5, 7]$, $[-\frac{5}{2}, \frac{3}{2}; \quad 2, -1]$, não existe, $[1, -\frac{2}{3}; \quad 2, -\frac{5}{3}]$

2.55 $[1, 1, -1; \quad 2, -5, 3; \quad -1, 2, -1]$, $[1, 1, 0; \quad -1, -3, 1; \quad -1, -4, 1]$

2.56 $A = [1, 2; \quad 1, 2]$, $B = [0, 0; \quad 1, 1]$, $C = [2, 2; \quad 0, 0]$

2.57 $A = [1, 2; \quad 0, 3]$; $B = [4, 3; \quad 3, 0]$

2.58 (c) *Sugestão*: Utilize o Problema 2.48

2.59 (a) $AB = \mathrm{diag}(2, -10, 0)$, $A^2 = \mathrm{diag}(1, 4, 9)$, $B^2 = \mathrm{diag}(4, 25, 0)$;

(b) $f(A) = \mathrm{diag}(2, 9, -6)$; (c) $A^{-1} = \mathrm{diag}(1, \frac{1}{2}, -\frac{1}{3})$, não existe C^{-1}.

2.60 (a) $[1, 2n; \quad 0, 1]$, (b) $[1, n, \frac{1}{2}n(n-1); \quad 0, 1, n; \quad 0, 0, 1]$

2.61 (a) $[2, 3; \quad 0, 5]$, $[-2, -3; \quad 0, -5]$, $[2, -7; \quad 0, -5]$, $[-2, 7; \quad 0, 5]$, (b) Não há.

2.62 (a) $k = 2$, (b) $k = -5$, (c) Não há.

2.63 $[1, 0; \quad 2, 3]$

2.64 $[1, 2, 1; \quad 0, 3, 1; \quad 0, 0, 2]$

2.65 Para uma matriz ser triangular superior, todas as entradas abaixo da diagonal devem ser 0 e para ser, também, invertível, todos os elementos diagonais devem ser 1.

(a) $8\ (2^n)$, (b) $2^6\ (2^{n(n+1)/2})$, (c) $2^3\ (2^{n(n-1)/2})$

2.67 (a) $A = [1, 1; \quad 0, 0]$, $B = [1, 2; \quad 3, 4]$, $C = [4, 6; \quad 0, 0]$

2.68 (a) $x = 4, y = 1, z = 3$; (b) $x = 0, y = -6$, z qualquer número real.

2.69 (c) *Sugestão*: Utilize $B = \frac{1}{2}(A + A^T)$ e $C = \frac{1}{2}(A - A^T)$.

2.70 $B = [4, 3; \quad 3, 3]$, $C = [0, 2; \quad -2, 0]$

2.72 (a) $[\frac{3}{5}, -\frac{4}{5}; \quad \frac{4}{5}, \frac{3}{5}]$, (b) $[1/\sqrt{5}, 2/\sqrt{5}; \quad 2/\sqrt{5}, -1/\sqrt{5}]$

2.73 (a) $[1/\sqrt{14}, 2/\sqrt{14}, 3/\sqrt{14};\quad 0, -2/\sqrt{13}, 3/\sqrt{13};\quad 12/\sqrt{157}, -3/\sqrt{157}, -2/\sqrt{157}]$
(b) $[1/\sqrt{11}, 3/\sqrt{11}, 1/\sqrt{11};\quad 1/\sqrt{2}, 0, -1/\sqrt{2};\quad 3/\sqrt{22}, -2/\sqrt{22}, 3/\sqrt{22}]$

2.75 A e C

2.76 $x = 3, y = 0, z = 3$

2.78 (c) *Sugestão*: Utilize $B = \frac{1}{2}(A + A^H)$ e $C = \frac{1}{2}(A - A^H)$.

2.79 As três.

2.81 A

2.82 (a) $UV = \text{diag}([7, 6;\quad 17, 10];\ [-1, 9;\quad 7, -5])$; (b) Não (c) Sim

2.83 Particione A com uma linha tracejada entre as primeira e segunda linhas (colunas); particione B com uma linha tracejada entre as segunda e terceira linhas (colunas) e entre as quarta e quinta linhas (colunas); não há como particionar C para torná-la uma matriz quadrada em blocos.

2.84 (a) $M^2 = \text{diag}([4],\ [9, 8;\quad 4, 9],\quad [9])$,
$M^3 = \text{diag}([8],\ [25, 44;\quad 22, 25],\ [27])$
(b) $M^2 = \text{diag}([3, 4;\quad 8, 11],\quad [9, 12;\quad 24, 33])$
$M^3 = \text{diag}([11, 15;\quad 30, 41],\quad [57, 78;\quad 156, 213])$

2.85 (a) $\text{diag}([7],\quad [8, 24;\quad 12, 8],\ [16])$, (b) $\text{diag}([2, 8;\quad 16, 181],\quad [8, 20;\quad 40, 48])$

Sistemas de Equações Lineares

3.1 INTRODUÇÃO

Os sistemas de equações lineares desempenham um papel importante e motivador na disciplina de Álgebra Linear. Na verdade, muitos problemas de Álgebra Linear se reduzem a encontrar uma solução de um sistema de equações lineares. Assim, as técnicas introduzidas neste capítulo serão aplicáveis às ideias abstratas que serão introduzidas mais tarde. Além disso, alguns dos resultados abstratos nos proporcionam uma nova perspectiva da estrutura e das propriedades de sistemas de equações lineares.

Todos os nossos sistemas de equações lineares contêm escalares, tanto como coeficientes quanto como constantes, e esses escalares podem ser elementos de um corpo K qualquer. Não há praticamente perda de generalidade alguma se o leitor supuser que todos os nossos escalares sejam números reais, isto é, elementos do corpo **R** dos reais.

3.2 DEFINIÇÕES BÁSICAS, SOLUÇÕES

Nesta seção apresentamos as definições básicas relativas às soluções de sistemas de equações lineares. Os algoritmos que utilizamos para determinar essas soluções só serão tratados adiante.

Equação linear e soluções

Uma *equação linear* nas incógnitas $x_1, x_2, ..., x_n$ é uma equação que pode ser colocada na seguinte *forma padrão*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad (3.1)$$

em que $a_1, a_2, ..., a_n$ e b são constantes. Dizemos que a constante a_k é o *coeficiente* de x_k e b é o *termo constante* da equação.

Uma solução da equação linear (3.1) é uma lista de valores para as incógnitas ou, equivalentemente, um vetor u de K^n, digamos,

$$x_1 = k_1, \quad x_2 = k_2, \quad \ldots, \quad x_n = k_n \qquad \text{ou} \qquad u = (k_1, k_2, \ldots, k_n)$$

tal que seja verdadeira a afirmação seguinte (obtida substituindo x_i por k_i na equação).

$$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$$

Nesse caso, dizemos que *u satisfaz* a equação.

OBSERVAÇÃO Na Equação (3.1), implicitamente, estamos supondo que há uma ordenação das incógnitas. Para evitar o uso de índices, costumamos usar x, y para duas incógnitas, x, y, z para três incógnitas e x, y, z, t para quatro incógnitas, que sempre ordenamos dessa forma.

Exemplo 3.1 Considere a equação linear nas três incógnitas x, y, z a seguir.

$$x + 2y - 3z = 6$$

Vemos que $x = 5$, $y = 2$, $z = 1$ ou, equivalentemente, o vetor $u = (5, 2, 1)$, é uma solução da equação. Isto é,

$$5 + 2(2) - 3(1) = 6 \quad \text{ou} \quad 5 + 4 - 3 = 6 \quad \text{ou} \quad 6 = 6$$

No entanto, $w = (1, 2, 3)$ não é uma solução, pois, substituindo esses valores, não obtemos uma afirmativa verdadeira.

$$1 + 2(2) - 3(3) = 6 \quad \text{ou} \quad 1 + 4 - 9 = 6 \quad \text{ou} \quad -4 = 6$$

Sistema de equações lineares

Um sistema de equações lineares é uma lista de equações com as mesmas incógnitas. Em particular, um sistema de m equações lineares $L_1, L_2, \ldots, L_m$ nas n incógnitas $x_1, x_2, \ldots, x_n$ pode ser colocado na *forma padrão*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3.2}$$

em que os a_{ij} e os b_i são constantes. O número a_{ij} é o *coeficiente* da incógnita x_j na equação L_i, e o número b_i é o *termo constante* dessa equação.

Dizemos que o sistema (3.2) é um sistema $m \times n$ (que lemos m por n). O sistema é dito *quadrado* se $m = n$, ou seja, se o número m de equações for igual ao número n de incógnitas.

Dizemos que o sistema (3.2) é *homogêneo* se todos os seus termos constantes forem nulos, isto é, se $b_1 = 0$, $b_2 = 0, \ldots, b_m = 0$. Caso contrário, o sistema é dito *não homogêneo*.

Uma *solução* (ou *solução particular*) do sistema (3.2) é uma lista de valores para as incógnitas ou, de modo equivalente, um vetor u de K^n, que é uma solução de cada uma das equações do sistema. O conjunto de todas as soluções do sistema é denominado *conjunto solução* ou *solução geral* do sistema.

Exemplo 3.2 Considere o sistema de equações lineares a seguir.

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= 5 \\ 2x_1 + 3x_2 + x_3 - 2x_4 &= 1 \\ x_1 + 2x_2 - 5x_3 + 4x_4 &= 3 \end{aligned}$$

Por ter três equações com quatro incógnitas, o sistema é 3×4. Decida se (a) $u = (-8, 6, 1, 1)$ e (b) $v = (-10, 5, 1, 2)$ são soluções do sistema.

(a) Substituímos os valores de u em cada equação, obtendo

$$\begin{array}{rcrcr} -8 + 6 + 4(1) + 3(1) = 5 & \text{ou} & -8 + 6 + 4 + 3 = 5 & \text{ou} & 5 = 5 \\ 2(-8) + 3(6) + 1 - 2(1) = 1 & \text{ou} & -16 + 18 + 1 - 2 = 1 & \text{ou} & 1 = 1 \\ -8 + 2(6) - 5(1) + 4(1) = 3 & \text{ou} & -8 + 12 - 5 + 4 = 3 & \text{ou} & 3 = 3 \end{array}$$

Sim, u é uma solução do sistema, pois é uma solução de cada equação.

(b) Substituímos os valores de u em cada equação, obtendo

$$\begin{array}{rcrcr} -10 + 5 + 4(1) + 3(2) = 5 & \text{ou} & -10 + 5 + 4 + 6 = 5 & \text{ou} & 5 = 5 \\ 2(-10) + 3(5) + 1 - 2(2) = 1 & \text{ou} & -20 + 15 + 1 - 4 = 1 & \text{ou} & -8 = 1 \end{array}$$

Não, v não é uma solução do sistema, pois não é uma solução da segunda equação. (Nem precisamos substituir v na terceira equação.)

Dizemos que o sistema (3.2) é *consistente* se tiver uma ou mais soluções; se não possuir solução alguma, dizemos que o sistema é *inconsistente*. Quando o corpo K de escalares é infinito, como, por exemplo, no caso em que K é o corpo real $\mathbf{R}$ ou o corpo complexo $\mathbf{C}$, vale o seguinte resultado importante.

Teorema 3.1 Seja K um corpo infinito. Então qualquer sistema $\mathscr{L}$ de equações lineares tem (i) uma solução, (ii) nenhuma solução, ou (iii) uma infinidade de soluções.

Esta situação é mostrada na Figura 3-1. Os três casos possuem uma interpretação geométrica se o sistema $\mathscr{L}$ consistir em duas equações com duas incógnitas (Seção 3.4).

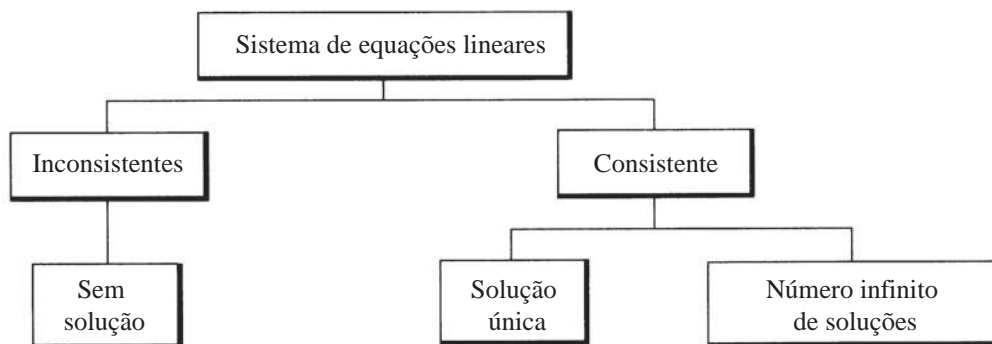



Figura 3-1

Matriz de coeficientes e matriz aumentada de um sistema

Voltemos ao sistema geral (3.2) de m equações com n incógnitas. Associamos a esse sistema duas matrizes, como segue.

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Dizemos que a primeira matriz M é a *matriz aumentada* do sistema e a segunda matriz A é a *matriz de coeficientes* do sistema.

A matriz de coeficientes é, simplesmente, a matriz dos coeficientes do sistema, sendo igual à matriz aumentada M do sistema sem a última coluna de termos constantes. É comum usar a notação $M = [A, B]$ para enfatizar as duas partes da matriz M, em que B denota o vetor coluna dos termos constantes. A matriz aumentada M e a de coeficientes A do sistema estudado no Exemplo 3.2 são

$$M = \begin{bmatrix} 1 & 1 & 4 & 3 & 5 \\ 2 & 3 & 1 & -2 & 1 \\ 1 & 2 & -5 & 4 & 3 \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} 1 & 1 & 4 & 3 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & -5 & 4 \end{bmatrix}$$

Como era de se esperar, A consiste em todas as colunas de M, exceto a última, que é a coluna dos termos constantes.

Claramente, um sistema de equações lineares fica completamente determinado por sua matriz aumentada M e vice-versa. Mais especificamente, cada linha de M corresponde a uma equação do sistema e cada coluna de M corresponde aos coeficientes de uma incógnita, exceto a última coluna, que corresponde aos termos constantes do sistema.

Equações lineares degeneradas

Uma equação linear é dita *degenerada* se todos seus coeficientes são nulos, isto é, se ela tem a forma

$$0x_1 + 0x_2 + \cdots + 0x_n = b \tag{3.3}$$

A solução de uma dessas equações depende apenas do valor da constante b. Mais precisamente,

(i) se $b \neq 0$, então a equação não possui solução e
(ii) se $b = 0$, então cada vetor $u = (k_1, k_2, ..., k_n)$ de K^n é uma solução.

Vale o resultado seguinte.

Teorema 3.2 Seja $\mathscr{L}$ um sistema de equações lineares que contém uma equação degenerada L com termo constante b.

(i) Se $b \neq 0$, então o sistema $\mathscr{L}$ não possui solução.
(ii) Se $b = 0$, então L pode ser retirada do sistema sem alterar seu conjunto solução.

A parte (i) decorre da observação que a equação degenerada não possui solução, por isso o sistema também não possui solução. A parte (ii) decorre da observação que cada elemento de K^n é solução da equação degenerada.

Incógnita líder numa equação linear não degenerada

Seja, agora, L uma equação linear não degenerada. Isso significa que um ou mais dos coeficientes de L não são nulos. A primeira incógnita de L com coeficiente não nulo é denominada *incógnita líder* de L. Por exemplo, x_3 e y são as incógnitas líderes, respectivamente, das equações

$$0x_1 + 0x_2 + 5x_3 + 6x_4 + 0x_5 + 8x_6 = 7 \qquad \text{e} \qquad 0x + 2y - 4z = 5$$

É costume omitir os termos com coeficientes nulos, portanto essas equações seriam escritas como

$$5x_3 + 6x_4 + 8x_6 = 7 \qquad \text{e} \qquad 2y - 4z = 5$$

Nesse caso, a incógnita líder é a que aparece em primeiro lugar.

3.3 SISTEMAS EQUIVALENTES, OPERAÇÕES ELEMENTARES

Voltemos ao sistema geral (3.2) de m equações com n incógnitas. Seja L a equação linear obtida multiplicando as m equações por constantes $c_1, c_2, ..., c_m$, respectivamente, e depois somando as equações resultantes. Mais precisamente, seja L a equação linear

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m$$

Dizemos que L é uma *combinação linear* das equações do sistema. É fácil mostrar (Problema 3.43) que qualquer solução do sistema (3.2) é também uma solução da combinação linear L.

Exemplo 3.3 Sejam L_1, L_2 e L_3, respectivamente, as três equações do Exemplo 3.2. Seja L a equação obtida multiplicando L_1, L_2 e L_3 por 3, –2 e 4 e somando, ou seja,

$$\begin{array}{rrcl}
3L_1: & 3x_1 + 3x_2 + 12x_3 + 9x_4 & = & 15 \\
-2L_2: & -4x_1 - 6x_2 - 2x_3 + 4x_4 & = & -2 \\
4L_1: & 4x_1 + 8x_2 - 20x_3 + 16x_4 & = & 12 \\
\hline
\text{(Soma) } L: & 3x_1 + 5x_2 - 10x_3 + 29x_4 & = & 25
\end{array}$$

Então L é uma combinação linear de L_1, L_2 e L_3. Como se esperava, a solução $u = (-8, 6, 1, 1)$ do sistema também é uma solução de L, ou seja, substituindo u em L obtemos uma afirmação verdadeira, a saber,

$$3(-8) + 5(6) - 10(1) + 29(1) = 25 \qquad \text{ou} \qquad -24 + 30 - 10 + 29 = 25 \qquad \text{ou} \qquad 9 = 9$$

Vale o teorema a seguir.

Teorema 3.3 Dois sistemas de equações lineares têm as mesmas soluções se, e só se, cada equação de cada sistema é uma combinação linear de equações do outro sistema.

Dois sistemas de equações lineares são ditos *equivalentes* se possuírem as mesmas soluções. Na próxima subseção mostramos uma maneira de obter sistemas equivalentes de equações lineares.

Operações elementares

As seguintes operações num sistema de equações lineares $L_1, L_2, ..., L_m$ são ditas *operações elementares*.

[E_1] Trocar duas equações de posição. Indicamos que as equações L_i e L_j foram trocadas de posição escrevendo

$$\text{“Trocar } L_i \text{ e } L_j \text{ entre si” ou “} L_i \longleftrightarrow L_j \text{”}$$

[E_2] Substituir uma equação por um múltiplo não nulo dela. Indicamos que a equação L_i foi substituída por kL_i (com $k \neq 0$) escrevendo

$$\text{“Substituir } L_i \text{ por } kL_i \text{” ou “} kL_i \to L_i \text{”}$$

[E_3] Substituir uma equação por um múltiplo de uma outra equação somado com ela mesma. Indicamos que a equação L_j foi substituída pela soma de kL_i e L_j escrevendo

$$\text{“Substituir } L_j \text{ por } kL_i + L_j \text{” ou “} kL_i + L_j \to L_j \text{”}$$

A seta $\to$ em [E_2] e [E_3] pode ser lida como “substitui”.

A propriedade principal dessas operações elementares está contemplada no teorema seguinte (demonstrado no Exercício 3.45).

Teorema 3.4 Suponha que um sistema $\mathscr{M}$ de equações lineares seja obtido de um sistema $\mathscr{L}$ de equações lineares por uma sequência finita de operações elementares. Então $\mathscr{M}$ e $\mathscr{L}$ têm as mesmas soluções.

OBSERVAÇÃO Às vezes (digamos, para evitar frações quando todos os escalares dados são inteiros) podemos aplicar [E_2] e [E_3] num único passo, isto é, podemos aplicar a operação seguinte.

[E] Substituir a equação L_j pela soma de kL_i e $k'L_j$ (com $k' \neq 0$), que indicamos por

$$\text{“Substituir } L_j \text{ por } kL_i + k'L_j \text{”} \qquad \text{ou} \qquad \text{“} kL_i + k'L_j \to L_j \text{”}$$

Enfatizamos que apenas a equação L_j é alterada nas operações [E_3] e [E].

> A eliminação gaussiana, nosso principal método para encontrar a solução de um dado sistema de equações lineares, consiste no uso das operações que acabamos de apresentar para transformar o dado sistema num sistema equivalente cuja solução possa ser obtida com facilidade.

Os detalhes da eliminação gaussiana são discutidos em seções subsequentes.

3.4 SISTEMAS QUADRADOS E PEQUENOS DE EQUAÇÕES LINEARES

Nesta seção consideramos os casos especiais de uma equação com uma incógnita e de duas equações com duas incógnitas. Como os conjuntos solução desses sistemas simples podem ser descritos geometricamente e suas propriedades motivam o caso geral, são tratados separadamente.

Equação linear com uma incógnita

O resultado a seguir é simples e será provado no Problema 3.4.

Teorema 3.5 Considere a equação linear $ax = b$.

(i) Se $a \neq 0$, então $x = b/a$ é a única solução de $ax = b$.
(ii) Se $a = 0$, mas $b \neq 0$, então $ax = b$ não possui solução.
(iii) Se $a = 0$ e $b = 0$, então cada escalar k é uma solução de $ax = b$.

Exemplo 3.4 Resolva (a) $4x - 1 = x + 6$, (b) $2x - 5 - x = x + 3$, (c) $4 + x - 3 = 2x + 1 - x$.

(a) Reescrevemos a equação na forma padrão, obtendo $3x = 7$. Então $x = \frac{7}{3}$ é a única solução [Teorema 3.5(i)].
(b) Reescrevemos a equação na forma padrão, obtendo $0x = 8$. A equação não possui solução [Teorema 3.5(ii)].
(c) Reescrevemos a equação na forma padrão, obtendo $0x = 0$. Então cada escalar k é uma solução [Teorema 3.5(iii)].

Sistema de duas equações lineares com duas incógnitas (sistemas 2×2)

Consideremos um sistema de duas equações lineares não degeneradas nas duas incógnitas x e y, que pode ser colocado na forma padrão

$$\begin{aligned} A_1x + B_1y &= C_1 \\ A_2x + B_2y &= C_2 \end{aligned} \tag{3.4}$$

Como ambas equações são não degeneradas, A_1 e B_1 não são simultaneamente nulos, nem o são A_2 e B_2.

A solução geral do sistema (3.4) é de um dos três tipos classificados na Figura 3-1. Se $\mathbf{R}$ for o corpo de escalares, então o gráfico de cada equação é uma reta no plano $\mathbf{R}^2$ e os três tipos podem ser interpretados geometricamente conforme indicado na Figura 3-2. Especificamente, temos os três tipos seguintes.

(1) *O sistema possui exatamente uma solução.*
Nesse caso, as duas retas se intersectam num ponto [Figura 3-2(a)]. Isto ocorre se as retas possuírem inclinações distintas ou, equivalentemente, se os coeficientes de x e y não forem proporcionais.

$$\frac{A_1}{A_2} \neq \frac{B_1}{B_2} \qquad \text{ou, equivalentemente,} \qquad A_1B_2 - A_2B_1 \neq 0$$

Por exemplo, na Figura 3-2(a), $1/3 \neq -1/2$.

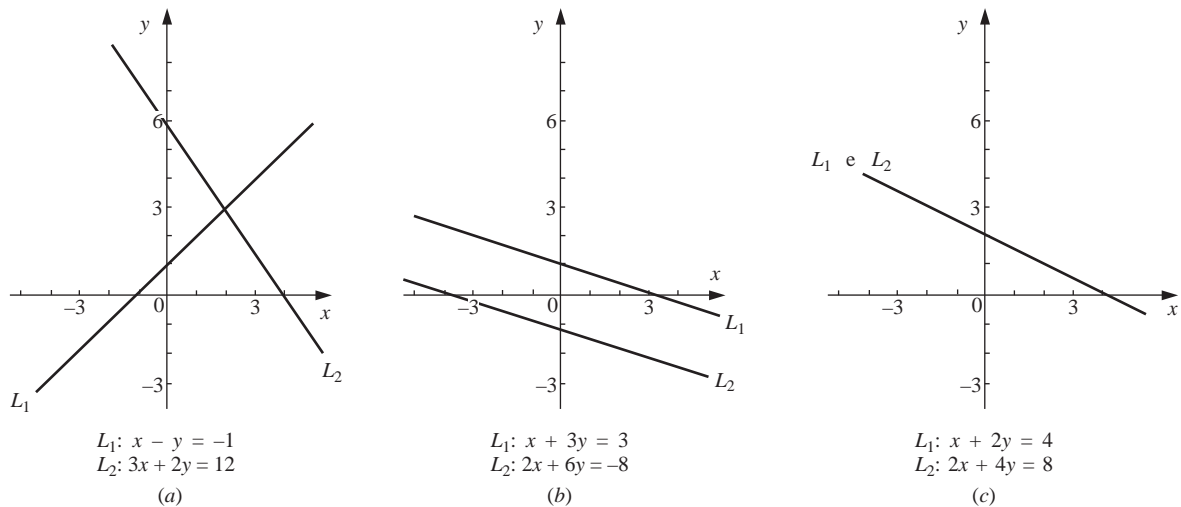



Figura 3-2

(2) *O sistema não possui solução*.
Nesse caso, as duas retas são paralelas e distintas [Figura 3-2(b)]. Isto ocorre se as retas possuírem a mesma inclinação, mas cruzarem o eixo y em pontos distintos, ou seja,

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$$

Por exemplo, na Figura 3-2(b), $1/2 = 3/6 \neq -3/8$.

(3) *O sistema possui uma infinidade soluções*.
Nesse caso, as duas retas coincidem [Figura 3-2(c)]. Isto ocorre se as retas possuírem a mesma inclinação e cruzarem o eixo y no mesmo ponto ou, equivalentemente, se os coeficientes e os termos constantes forem proporcionais.

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$$

Por exemplo, na Figura 3-2(c), $1/2 = 2/4 = 4/8$.

OBSERVAÇÃO A expressão seguinte e seu valor são denominados *determinante de ordem dois*.

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = A_1B_2 - A_2B_1$$

Os determinantes serão estudados no Capítulo 8. Assim, o sistema (3.4) possui solução única se, e só se, o determinante de seus coeficientes não é zero. (Mais adiante, mostraremos que essa afirmação é verdadeira para qualquer sistema quadrado de equações lineares.)

Algoritmo de eliminação

A solução do sistema (3.4) pode ser obtida pelo processo de eliminação, em que reduzimos o sistema a uma única equação com uma única incógnita. Supondo que o sistema possua uma única solução, esse algoritmo de eliminação tem duas partes.

Algoritmo 3.1 É dado um sistema de duas equações lineares não degeneradas L_1 e L_2 com duas incógnitas e uma única solução.

Parte A (Eliminação Para Frente) Multiplique cada equação por uma constante para que os coeficientes resultantes de uma das incógnitas sejam simétricos e, em seguida, some as duas equações, obtendo uma nova equação L com apenas uma incógnita.

Parte B (Substituição Para Trás) Resolva a nova equação L (de apenas uma incógnita), substitua esse valor da incógnita em uma das equações originais e resolva-a para obter o valor da outra incógnita.

A Parte A do Algoritmo 3.1 pode ser aplicada a qualquer sistema, mesmo que o sistema não possua solução única. Nesse caso, a nova equação L será degenerada e a parte B não poderá ser aplicada.

Exemplo 3.5 (Solução Única) Resolva o sistema

$$\begin{aligned} L_1&: \ 2x - 3y = -8 \\ L_2&: \ 3x + 4y = \ \ \ 5 \end{aligned}$$

A incógnita x é eliminada das equações formando a nova equação $L = -3L_1 + 2L_2$, ou seja, multiplicamos L_1 por -3 e L_2 por 2 e depois somamos as equações resultantes, como segue.

$$\begin{array}{rr} -3L_1: & -6x + 9y = 24 \\ 2L_2: & 6x + 8y = 10 \\ \hline \text{Soma}: & 17y = 34 \end{array}$$

Resolvendo a nova equação em y, obtemos $y = 2$. Substituímos $y = 2$ numa das equações originais, digamos, L_1, e resolvemos na outra variável, x, obtendo

$$2x - 3(2) = -8 \quad \text{ou} \quad 2x - 6 = 8 \quad \text{ou} \quad 2x = -2 \quad \text{ou} \quad x = -1$$

Assim, $x = -1$, $y = 2$ ou, então, o par $u = (-1, 2)$ é a única solução do sistema. Como $2/3 \neq -3/4$, já estávamos esperando que a solução fosse única. [Geometricamente, as retas correspondentes às equações têm interseção no ponto (–1, 2).]

Exemplo 3.6 (Solução não única)

(a) Resolva o sistema

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = 5 \end{aligned}$$

Eliminamos x das equações multiplicando L_1 por 2 e, depois, somando com L_2, ou seja, formamos a nova equação $L = 2L_1 + L_2$. Isso fornece a equação degenerada

$$0x + 0y = 13$$

em que o termo constante $b = 13$ é não nulo. Assim, essa equação, e o sistema, não têm solução. Isto era de se esperar, pois $1/(-2) = -3/6 \neq 4/5$. [Geometricamente, as retas correspondentes às equações são paralelas e não coincidem.]

(b) Resolva o sistema

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = -8 \end{aligned}$$

Eliminamos x das equações multiplicando L_1 por 2 e, depois, somando com L_2, ou seja, formamos a nova equação $L = 2L_1 + L_2$. Isso fornece a equação degenerada

$$0x + 0y = 0$$

em que o termo constante é nulo. Assim, o sistema tem uma infinidade de soluções, que correspondem às soluções de cada uma das equações. Isso era de se esperar, pois $1/(-2) = -3/6 = 4/(-8)$. [Geometricamente, as retas correspondentes às equações coincidem.]

Para encontrar a solução geral, substituímos $y = a$ em L_1 para obter

$$x - 3a = 4 \quad \text{ou} \quad x = 3a + 4$$

Assim, a solução geral do sistema é

$$x - 3a = 4, y = a \quad \text{ou} \quad u = (3a + 4, a)$$

em que o *parâmetro a* é um escalar qualquer.

3.5 SISTEMAS EM FORMA TRIANGULAR E ESCALONADA

A eliminação gaussiana, que é o principal método de resolução de sistemas de equações lineares, é tratada na Seção 3.6. Aqui consideramos dois tipos simples de sistemas de equações lineares, a saber, os triangulares e, mais geralmente, os escalonados.

Forma triangular

Considere o sistema de equações lineares em *forma triangular* a seguir.

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 - 2x_4 &= 9 \\ 5x_2 - \ \ x_3 + 3x_4 &= 1 \\ 7x_3 - \ \ x_4 &= 3 \\ 2x_4 &= 8 \end{aligned}$$

Num sistema em forma triangular, a primeira incógnita x_1 é a incógnita líder da primeira equação, a segunda incógnita x_2 é a incógnita líder da segunda equação e assim por diante. Em particular, o sistema é quadrado e cada incógnita líder está *imediatamente* à direita da incógnita líder da equação precedente.

Um sistema triangular desses sempre possui solução única, que pode ser obtida por substituição para trás, como segue.

(1) Primeiro resolvemos a última equação, para obter $x_4 = 4$.
(2) Em seguida, substituímos esse valor $x_4 = 4$ na penúltima equação e resolvemos na penúltima variável x_3, como segue.

$$7x_3 - 4 = 3 \qquad \text{ou} \qquad 7x_3 = 7 \text{ ou } x_3 = 1$$

(3) Agora substituímos $x_3 = 1$ e $x_4 = 4$ na segunda equação e resolvemos na segunda variável x_2, como segue.

$$5x_2 - 1 + 12 = 1 \qquad \text{ou} \qquad 5x_2 + 11 = 1 \qquad \text{ou} \qquad 5x_2 = -10 \text{ ou } x_2 = -2$$

(4) Finalmente, substituímos $x_2 = -2$, $x_3 = 1$ e $x_4 = 4$ na primeira equação e resolvemos na primeira variável x_1, como segue.

$$2x_1 + 6 + 5 - 8 = 9 \qquad \text{ou} \qquad 2x_1 + 3 = 9 \qquad \text{ou} \qquad 2x_1 = 6 \text{ ou } x_1 = 3$$

Assim, $x_1 = 3$, $x_2 = -2$, $x_3 = 1$ e $x_4 = 4$ ou, equivalentemente, o vetor $u = (3, -2, 1, 4)$ é a única solução do sistema.

OBSERVAÇÃO Existe uma forma alternativa para a substituição para trás (que será utilizada quando resolvermos um sistema na forma matricial). A saber, depois de encontrar o valor da última incógnita, substituímos esse valor em todas as equações que a precedem antes de resolver a penúltima equação. Isso fornece um sistema triangular com uma equação e uma incógnita a menos. Por exemplo, no sistema triangular que acabamos de resolver, substituímos $x_4 = 4$ em todas as equações anteriores para obter o sistema triangular

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 17 \\ 5x_2 - \ \ x_3 &= -1 \\ 7x_3 &= 7 \end{aligned}$$

Então repetimos o processo usando a nova última equação e, assim, sucessivamente.

Forma escalonada, pivô e variáveis livres

Dizemos que o sistema de equações lineares seguinte está em *forma escalonada*.

$$\begin{aligned} 2x_1 + 6x_2 - x_3 + 4x_4 - 2x_5 &= 15 \\ x_3 + 2x_4 + 2x_5 &= 5 \\ 3x_4 - 9x_5 &= 6 \end{aligned}$$

Num sistema em forma escalonada, nenhuma equação é degenerada, e a incógnita líder de cada equação a partir da segunda está à direita da incógnita líder da equação precedente. As incógnitas líderes das equações do sistema, no exemplo, x_1, x_3 e x_4, são denominadas variáveis *pivô* e as outras incógnitas, x_2 e x_5 no exemplo, são denominadas variáveis *livres*.

Em geral, um *sistema escalonado* ou um *sistema em forma escalonada* é da forma seguinte.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= b_r \end{aligned} \tag{3.5}$$

onde $1 < j_2 < \cdots < j_r$ e $a_{11}, a_{2j_2}, \ldots, a_{rj_r}$ são não nulos. As variáveis pivô são $x_1, x_{j_2}, \ldots, x_{j_r}$. Observe que $r \leq n$.

O conjunto solução de um sistema escalonado é descrito no teorema seguinte (demonstrado no Problema 3.10).

Teorema 3.6 Considere um sistema de equações lineares em forma escalonada, digamos, de r equações com n incógnitas. Há dois casos.

(i) $r = n$, ou seja, há o mesmo número de equações e incógnitas (forma triangular). Nesse caso, o sistema tem uma única solução.

(ii) $r < n$, ou seja, há mais incógnitas do que equações. Nesse caso, podemos atribuir livremente valores às $n - r$ variáveis livres e resolver de modo único nas r variáveis pivô, obtendo uma solução do sistema.

Suponha que um sistema escalonado contenha mais incógnitas do que equações. Supondo que o corpo K seja infinito, o sistema possui uma infinidade de soluções, porque a cada uma das $n - r$ variáveis livres podemos atribuir qualquer valor escalar.

A solução geral de um sistema com variáveis livres pode ser descrita de uma de duas maneiras equivalentes, que ilustramos usando o sistema escalonado de $r = 3$ equações e $n = 5$ incógnitas dado acima. Uma descrição é dita "forma paramétrica" da solução e a outra "forma de variáveis livres".

Forma paramétrica

Atribuímos valores arbitrários, denominados *parâmetros*, às variáveis livres x_2 e x_5, digamos, $x_2 = a$ e $x_5 = b$, e utilizamos a substituição para trás para obter os valores das variáveis pivô x_1, x_3 e x_4 em termos dos parâmetros a e b, como segue.

(1) Substituímos $x_5 = b$ na última equação e resolvemos em x_4.

$$3x_4 - 9b = 6 \quad \text{ou} \quad 3x_4 = 6 + 9b \quad \text{ou} \quad x_4 = 2 + 3b$$

(2) Substituímos $x_4 = 2 + 3b$ e $x_5 = b$ na segunda equação e resolvemos em x_3.

$$x_3 + 2(2 + 3b) + 2b = 5 \quad \text{ou} \quad x_3 + 4 + 8b = 5 \quad \text{ou} \quad x_3 = 1 - 8b$$

(3) Substituímos $x_2 = a$, $x_3 = 1 - 8b$, $x_4 = 2 + 3b$ e $x_5 = b$ na primeira equação e resolvemos em x_1.

$$2x_1 + 6a - (1 - 8b) + 4(2 + 3b) - 2b = 15 \quad \text{ou} \quad x_1 = 4 - 3a - 9b$$

Assim, a solução geral em *forma paramétrica* é

$$x_1 = 4 - 3a - 9b, \quad x_2 = a, \quad x_3 = 1 - 8b, \quad x_4 = 2 + 3b, \quad x_5 = b$$

ou, equivalentemente, $v = (4 - 3a - 9b, a, 1 - 8b, 2 + 3b, b)$, em que a e b são escalares arbitrários.

Forma de variáveis livres

Utilizamos a substituição para trás para resolver nas variáveis pivô x_1, x_3 e x_4 diretamente em termos das variáveis livres pivô x_2 e x_5. Ou seja, a última equação fornece $x_4 = 2 + 3x_5$, substituindo na segunda equação fornece $x_3 = 1 - 8x_5$ e substituindo na primeira equação fornece $x_1 = 4 - 3x_2 - 9x_5$. Assim, obtemos

$$x_1 = 4 - 3x_2 - 9x_5, \quad x_2 = \text{variável livre}, \quad x_3 = 1 - 8x_5, \quad x_4 = 2 + 3x_5, \quad x_5 = \text{variável livre}$$

ou, equivalentemente,

$$v = (4 - 3x_2 - 9x_5, x_2, 1 - 8x_5, 2 + 3x_5, x_5),$$

é a *forma de variáveis livres* da solução geral do sistema.

Enfatizamos que não há diferença alguma entre as duas formas da solução geral, sendo a utilização de uma ou outra simplesmente uma questão de preferência pessoal.

OBSERVAÇÃO Uma solução particular do sistema estudado pode ser encontrada atribuindo quaisquer valores às variáveis livres e, então, resolvendo nas variáveis pivô com substituição para trás. Por exemplo, tomando $x_2 = 1$ e $x_5 = 1$, obtemos

$$x_4 = 2 + 3 = 5, \qquad x_3 = 1 - 8 = -7, \qquad x_1 = 4 - 3 - 9 = -8$$

Assim, $u = (-8, 1, 7, 5, 1)$ é a solução particular correspondente a $x_2 = 1$ e $x_5 = 1$.

3.6 ELIMINAÇÃO GAUSSIANA

O método principal de resolução do sistema de equações lineares (3.2) gerais é denominado *eliminação gaussiana* e consiste, essencialmente, em duas partes.

Parte A (Eliminação Para Frente) Redução passo a passo do sistema, levando ou a uma equação degenerada sem solução (indicando que o sistema não tem solução) ou a um sistema equivalente mais simples em forma triangular ou escalonada.

Parte B (Substituição Para Trás) Substituição para trás passo a passo para obter a solução do sistema mais simples.

A parte B já foi discutida nas seções precedentes, portanto, basta dar o algoritmo para a Parte A, como segue.

Algoritmo 3.2 (para Parte A) É dado um sistema $m \times n$ de equações lineares (3.2).

Etapa de eliminação Encontre a primeira incógnita do sistema com um coeficiente não nulo (que agora, obrigatoriamente, é x_1).

(a) Faça com que $a_{11} \neq 0$, ou seja, se necessário, troque de lugar as equações para que a primeira incógnita x_1 apareça com um coeficiente não nulo na primeira equação.

(b) Use a_{11} *como um pivô* para eliminar x_1 de todas as equações, exceto a primeira. Em outras palavras, para cada $i > 1$,

(1) denote $m = -a_{i1}/a_{11}$; (2) substitua L_i por $mL_1 + L_i$

Agora, o sistema tem a forma seguinte,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots& \\ a_{mj_2}x_{j_2} + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

em que x_1 aparece somente na primeira equação, $a_{11} \neq 0$ e x_{j_2} denota a primeira incógnita com um coeficiente não nulo nas equações depois da primeira.

(c) Examine cada nova equação L.

(1) Se L tiver a forma $0x_1 + 0x_2 + \cdots + 0x_n = b$, com $b \neq 0$, então

PARE

O sistema é *inconsistente* e não possui solução.

(2) Se L tiver a forma $0x_1 + 0x_2 + \cdots + 0x_n = 0$, ou se L for um múltiplo de alguma outra equação, então retire L do sistema.

Etapa de recursão Repita a Etapa de Eliminação para cada novo subsistema "menor" formado por todas as equações do sistema, exceto a primeira equação.

Resultado Ao final, o sistema está reduzido à forma triangular ou escalonada, a menos que tenha sido encontrada alguma equação degenerada sem solução, indicando que o sistema é inconsistente.

As observações a seguir são relativas à Etapa de Eliminação do Algoritmo 3.2.

(1) Dizemos que o número m dado em (b) é um *multiplicador*.

$$m = -\frac{a_{i1}}{a_{11}} = -\frac{\text{coeficiente a ser eliminado}}{\text{pivô}}$$

(2) Alternativamente, poderíamos também aplicar a operação seguinte em (b).

Substituir L_i por $-a_{i1}L_1 + a_{11}L_i$

Isto evitaria o aparecimento de frações se todos os escalares, inicialmente, forem inteiros.

Exemplo de eliminação gaussiana

Exemplifiquemos em detalhes a eliminação gaussiana usando o sistema de equações lineares seguinte.

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2x - 4y - 3z &= 8 \\ L_3&: & -3x + 6y + 8z &= -5 \end{aligned}$$

Parte A Usamos o coeficiente 1 de x na primeira equação L_1 como o pivô para eliminar x da segunda equação L_2 e da terceira equação L_3. Isso é feito como segue.

(1) Multiplicamos L_1 pelo multiplicador $m = -2$ e somamos com L_2, ou seja, "substitua L_2 por $-2L_1 + L_2$".

(2) Multiplicamos L_1 pelo multiplicador $m = 3$ e somamos com L_3, ou seja, "substitua L_3 por $3L_1 + L_3$".

Com esses passos obtemos

$$\begin{aligned} (-2)L_1&: & -2x + 6y + 4z &= -12 \\ L_2&: & 2x - 4y - 3z &= 8 \\ \hline \text{Nova } L_2&: & 2y + z &= -4 \end{aligned} \qquad \begin{aligned} 3L_1&: & 3x - 9y - 6z &= 18 \\ L_3&: & -3x + 6y + 8z &= -5 \\ \hline \text{Nova } L_3&: & -3y + 2z &= 13 \end{aligned}$$

Assim, o sistema original foi substituído pelo sistema seguinte.

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & -3y + 2z &= 13 \end{aligned}$$

(Observe que as novas equações L_2 e L_3 formam um subsistema com uma equação a menos e uma incógnita a menos do que o sistema original.)

Agora usamos o coeficiente 2 de y na (nova) segunda equação L_2 como o pivô para eliminar y da (nova) terceira equação L_3. Isso é feito como segue.

(3) Multiplicamos L_2 pelo multiplicador $m = \frac{3}{2}$ e somamos com L_3, ou seja, "substitua L_3 por $\frac{3}{2}L_2 + L_3$". (Alternativamente, "substitua L_3 por $3L_2 + 2L_3$", que evita frações.)

Com esse passo obtemos

$$\begin{aligned} \tfrac{3}{2}L_2&: & 3y + \tfrac{3}{2}z &= -6 \\ L_3&: & -3y + 2z &= 13 \\ \hline \text{Nova } L_3&: & \tfrac{7}{2}z &= 7 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} 3L_2&: & 6y + 3z &= -12 \\ 2L_3&: & -6y + 4z &= 26 \\ \hline \text{Nova } L_3&: & 7z &= 14 \end{aligned}$$

Assim, nosso sistema original foi substituído pelo sistema seguinte.

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & 7z &= 14 \quad (\text{ou } \tfrac{7}{2}z = 7) \end{aligned}$$

Agora o sistema está em forma triangular, portanto, completamos a Parte A.

Parte B Os valores das incógnitas são obtidos na ordem inversa, z, y, x, por substituição para trás, como segue.

(1) Resolvemos z em L_3 para obter $z = 2$.

(2) Substituímos $z = 2$ em L_2 e resolvemos y para obter $y = -3$.

(3) Substituímos $y = -3$ e $z = 2$ em L_1 e resolvemos x para obter $x = 1$.

Assim, a solução do sistema triangular e, portanto, do sistema original é

$$x = 1, y = -3, z = 2 \qquad \text{ou, equivalentemente,} \qquad u = (1, -3, 2).$$

Formato condensado

O algoritmo de eliminação gaussiana envolve reescrever os sistemas de equações lineares. Às vezes, podemos evitar a cópia excessiva de algumas equações utilizando o "formato condensado". Apresentamos esse formato para a solução do sistema precedente.

Número	Equação	Operação
(1)	$x - 3y - 2z = 6$	
(2)	$2x - 4y - 3z = 8$	
(3)	$-3x + 6y + 8z = -5$	
$(2')$	$2y + z = -4$	Substitua L_2 por $-2L_1 + L_2$
$(3')$	$-3y + 2z = 13$	Substitua L_3 por $3L_1 + L_3$
$(3'')$	$7z = 14$	Substitua L_3 por $3L_2 + 2L_3$

Ou seja, inicialmente escrevemos o número de cada uma das equações originais. Ao aplicar a eliminação gaussiana ao sistema, apenas escrevemos as novas equações e as identificamos com o mesmo número da equação original correspondente, mas com um apóstrofo. (Após cada nova equação, indicamos, apenas por razões didáticas, a operação elementar que forneceu a nova equação.)

O sistema em forma triangular consiste nas equações (1), $(2')$ e $(3'')$, que são os números com maior número de apóstrofos. A substituição para trás nessas equações novamente fornece $x = 1$, $y = -3$, $z = 2$.

OBSERVAÇÃO Se duas equações necessitam ser trocadas de lugar, digamos, para obter um coeficiente não nulo como um pivô, então isso pode ser facilmente realizado nesse formato, simplesmente renumerando as duas equações, em vez de trocar suas posições.

Exemplo 3.7 Resolva o sistema

$$\begin{aligned} x + 2y - 3z &= 1 \\ 2x + 5y - 8z &= 4 \\ 3x + 8y - 13z &= 7 \end{aligned}$$

Resolvemos esse sistema com eliminação gaussiana.

Parte A (Eliminação Para Frente) Usamos o coeficiente 1 de x na primeira equação L_1 como o pivô para eliminar x da segunda equação L_2 e da terceira equação L_3. Isso é feito como segue.

(1) Multiplicamos L_1 pelo multiplicador $m = -2$ e somamos com L_2, ou seja, "substitua L_2 por $-2L_1 + L_2$".
(2) Multiplicamos L_1 pelo multiplicador $m = -3$ e somamos com L_3, ou seja, "substitua L_3 por $-3L_1 + L_3$".

Esses dois passos fornecem

$$\begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \\ 2y - 4z &= 4 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \end{aligned}$$

(A terceira equação foi suprimida, por ser um múltiplo da segunda equação.) Agora o sistema está em forma escalonada com variável livre z.

Parte B (Eliminação Para Trás) Para obter a solução geral, tomamos a variável livre $z = a$ e resolvemos em x e y por substituição para trás. Substituindo $z = a$ na segunda equação, obtemos $y = 2 + 2a$ e, substituindo $z = a$ e $y = 2 + 2a$ na primeira equação, obtemos

$$x + 2(2 + 2a) - 3a = 1 \qquad \text{ou} \qquad x + 4 + 4a - 3a = 1 \qquad \text{ou} \qquad x = -3 - a$$

Assim, a solução geral, em que a é um parâmetro, é

$$x = -3 - a, \qquad y = 2 + 2a, \qquad z = a \qquad \text{ou} \qquad u = (-3 - a, 2 + 2a, a).$$

Exemplo 3.8 Resolva o sistema

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ 2x_1 + 8x_2 - x_3 + 9x_4 &= 9 \\ 3x_1 + 5x_2 - 12x_3 + 17x_4 &= 7 \end{aligned}$$

Utilizamos eliminação gaussiana.

Parte A (Eliminação Para Frente) Usamos o coeficiente 1 de x_1 na primeira equação L_1 como o pivô para eliminar x_1 da segunda equação L_2 e da terceira equação L_3. Isso é feito com as operações seguintes.

(1) "Substitua L_2 por $-2L_1 + L_2$" e (2) "Substitua L_3 por $-3L_1 + L_3$".

Essas operações fornecem

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ 2x_2 + 3x_3 - x_4 &= 1 \\ -4x_2 - 6x_3 + 2x_4 &= -5 \end{aligned}$$

Agora utilizamos o coeficiente 2 de x_2 na segunda equação L_2 como o pivô e o multiplicador $m = 2$ para eliminar x_2 da terceira equação L_3. Isso é feito com a operação "substitua L_3 por $2L_2 + L_3$", que, então fornece a equação degenerada

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -3$$

Essa equação e, portanto, o sistema original, não possui solução.

NÃO CONTINUE

OBSERVAÇÃO 1 Como nos exemplos precedentes, a Parte A da eliminação nos diz se o sistema possui ou não uma solução, ou seja, se o sistema é ou não consistente. Assim, a Parte B nunca precisa ser aplicada quando o sistema não possuir solução.

OBSERVAÇÃO 2 Se um sistema de equações lineares tiver mais do que quatro equações e quatro incógnitas, pode ser mais conveniente usar o formato matricial para resolvê-lo. Esse formato matricial será discutido adiante.

3.7 MATRIZES ESCALONADAS, FORMA CANÔNICA POR LINHAS, EQUIVALÊNCIA POR LINHAS

Uma maneira de resolver um sistema de equações lineares é trabalhar, não com o próprio sistema, mas com sua matriz aumentada M. Nesta seção introduzimos os conceitos matriciais necessários para isso. Esses conceitos, como as matrizes escalonadas e operações elementares por linhas, também têm interesse próprio.

Matrizes escalonadas

Dizemos que uma matriz A é *escalonada*, ou que está em *forma escalonada*, se valerem as duas condições seguintes (em que um *elemento líder não nulo* de uma linha de A significa o primeiro elemento não nulo dessa linha).

(1) Todas as linhas nulas, se existirem, estão agrupadas juntas nas linhas inferiores da matriz.
(2) Cada elemento líder não nulo de uma linha está à direita do elemento líder não nulo da linha precedente.

Ou seja, $A = [a_{ij}]$ é uma matriz escalonada se existirem entradas não nulas

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}, \qquad \text{com} \qquad j_1 < j_2 < \cdots < j_r$$

com a propriedade de que

$$a_{ij} = 0 \qquad \text{com} \quad \begin{cases} \text{(i) } i \le r, \quad j < j_i \\ \text{(ii) } i > r \end{cases}$$

As entradas $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$, que são os elementos líderes não nulos de suas respectivas linhas, são denominadas *pivôs* da matriz escalonada.

Exemplo 3.9 A matriz seguinte é uma matriz escalonada cujos pivôs estão circulados.

$$A = \begin{bmatrix} 0 & \textcircled{2} & 3 & 4 & 5 & 9 & 0 & 7 \\ 0 & 0 & 0 & \textcircled{3} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{5} & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{8} & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe que os pivôs estão nas colunas C_2, C_4, C_6 e C_7 e que cada um está à direita do que o precede. Usando a notação precedente,

$$a_{1j_1} = 2, \qquad a_{2j_2} = 3, \qquad a_{3j_3} = 5, \qquad a_{4j_4} = 8$$

com $j_1 = 2, j_2 = 4, j_3 = 6, j_4 = 7$. Aqui, $r = 4$.

Forma canônica por linhas

Uma matriz A é dita em *forma canônica por linhas* ou, então, em *forma escalonada reduzida por linhas*, se for uma matriz escalonada, ou seja, satisfaz as condições (1) e (2) precedentes e, também, as duas condições seguintes.

(3) Cada pivô (elemento líder não nulo) é igual a 1.
(4) Cada pivô é o único elemento não nulo em sua coluna.

A principal diferença entre uma matriz escalonada e uma matriz em forma canônica por linhas é que, numa matriz escalonada, todas as entradas abaixo do pivô são nulas [Propriedades (1) e (2)], mas numa matriz na forma canônica por linhas, cada pivô também deve ser igual a 1 [Propriedade (3)] e todas as entradas acima do pivô são nulas [Propriedade (4)].

As matrizes nula 0 e identidade I de qualquer tamanho são exemplos especiais importantes de matrizes em forma canônica por linhas.

Exemplo 3.10

As matrizes seguintes são matrizes escalonadas cujos pivôs estão circulados.

$$\begin{bmatrix} \textcircled{2} & 3 & 2 & 0 & 4 & 5 & -6 \\ 0 & 0 & \textcircled{0} & \textcircled{1} & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{6} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} \textcircled{1} & 2 & 3 \\ 0 & 0 & \textcircled{1} \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \textcircled{1} & 3 & 0 & 0 & 4 \\ 0 & 0 & 0 & \textcircled{1} & 0 & -3 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 2 \end{bmatrix}$$

A terceira matriz também é um exemplo de uma matriz em forma canônica por linhas. A segunda matriz não está em forma canônica por linhas, por não satisfazer a Propriedade (4), já que possui uma entrada não nula acima do segundo pivô da terceira coluna. A primeira matriz não está em forma canônica por linhas, por não satisfazer nem a Propriedade (3) nem a (4), já que alguns pivôs não são iguais a 1 e há elementos não nulos acima dos pivôs.

Operações elementares com as linhas

Suponha que A seja uma matriz com linhas $R_1, R_2, ..., R_m$. As seguintes operações sobre A são denominadas *operações elementares sobre as linhas*.

[E_1] (Troca da Ordem de Linhas) Trocar de posição as linhas R_i e R_j. Isso pode ser escrito como segue.

$$\text{“Trocar } R_i \text{ e } R_j \text{ entre si”} \qquad \text{ou} \qquad \text{“}R_i \longleftrightarrow R_j\text{”}$$

[E_2] (Multiplicação de Linha por Escalar) Substituir a linha R_i por um múltiplo não nulo kR_i da linha R_i. Isso pode ser escrito como segue.

$$\text{“Substituir } R_i \text{ por } kR_i\,(k \neq 0) \qquad \text{ou} \qquad \text{“}kR_i \to R_i\text{”}$$

[E_3] (Soma de Linhas) Substituir a linha R_j pela soma de um múltiplo kR_i de uma linha R_i com a própria linha R_j. Isso pode ser escrito como segue.

$$\text{“Substituir } R_j \text{ por } kR_i + R_j\text{”} \qquad \text{ou} \qquad \text{“}kR_i + R_j \to R_j\text{”}$$

A seta $\to$ nas operações E_2 e E_3 pode ser lida como “substitui”.

Às vezes (digamos, para evitar frações quando todos os escalares dados são inteiros), podemos aplicar [E_2] e [E_3] num único passo, ou seja, podemos aplicar a operação seguinte.

[E] Substituir a linha R_j pela soma de um múltiplo kR_i de uma linha R_i com um múltiplo não nulo $k'R_j$ da própria linha R_j. Isso pode ser escrito como segue.

$$\text{“Substituir } R_j \text{ por } kR_i + k'R_j\ (k' \neq 0)\text{”} \qquad \text{ou} \qquad \text{“}kR_i + k'R_j \to R_j\text{”}$$

Enfatizamos que nas operações [E_3] e [E] somente trocamos a linha R_j.

Equivalência por linhas, posto de uma matriz

Dizemos que uma matriz A é *equivalente por linhas* a uma matriz B, o que é indicado por

$$A \sim B$$

se B puder ser obtida a partir de A por uma sequência de operações elementares com as linhas. Se B também for uma matriz escalonada, dizemos que B é uma *forma escalonada* de A.

Dois resultados básicos sobre a equivalência por linhas são os seguintes.

Teorema 3.7 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ duas matrizes escalonadas equivalentes por linhas com respectivas entradas pivô

$$a_{1j_1}, a_{2j_2}, \dots a_{rj_r} \qquad \text{e} \qquad b_{1k_1}, b_{2k_2}, \dots b_{sk_s}$$

Então A e B têm o mesmo número de linhas não nulas, ou seja, $r = s$, e as entradas pivô estão nas mesmas posições, ou seja, $j_1 = k_1, \quad j_2 = k_2, \quad \dots, \quad j_r = k_r$.

Teorema 3.8 Toda matriz A é equivalente por linhas a uma única matriz em forma canônica por linhas.

As demonstrações desses teoremas serão adiadas até o Capítulo 4. A matriz única no Teorema 3.8 é denominada *forma canônica por linhas* de A.

Usando esses dois teoremas, podemos dar nossa primeira definição de posto de uma matriz.

DEFINIÇÃO O *posto* de uma matriz A, denotado por pos(A), é igual ao número de pivôs numa forma escalonada de A.

O posto é uma propriedade muito importante de uma matriz e, dependendo do contexto em que utilizamos as matrizes, pode ser definido de várias maneiras diferentes. É claro que todas as definições sempre dão o mesmo número.

Na próxima seção apresentamos a forma matricial da eliminação gaussiana, que determina uma forma escalonada de qualquer matriz A (e, portanto, o posto de A) e também determina a forma canônica por linhas de A.

Pode ser demonstrado que a equivalência por linhas é uma *relação de equivalência*. Isto significa que

(1) $A \sim A$, para qualquer matriz A,
(2) se $A \sim B$, então $B \sim A$ e
(3) se $A \sim B$ e $B \sim C$, então $A \sim C$.

A propriedade (2) decorre do fato de que cada operação elementar com as linhas tem uma operação inversa do mesmo tipo, como segue.

(i) "Trocar R_i e R_j entre si" é sua própria inversa.
(ii) "Substituir R_i por kR_i" e "Substituir R_i por $(1/k)R_i$" são inversas.
(iii) "Substituir R_j por $kR_i + R_j$" e "Substituir R_j por $-kR_i + R_j$" são inversas.

Para a operação E há um resultado análogo (Problema 3.73).

3.8 ELIMINAÇÃO GAUSSIANA, FORMULAÇÃO MATRICIAL

Nesta seção, apresentamos dois algoritmos matriciais que têm os resultados seguintes.

(1) O algoritmo 3.3 transforma qualquer matriz A numa forma escalonada.
(2) O algoritmo 3.4 transforma a matriz escalonada em sua forma canônica por linhas.

Esses algoritmos, que usam as operações elementares com as linhas, são simplesmente reformulações da *eliminação gaussiana* aplicada a matrizes em vez de equações lineares. (O termo "reduzir por linhas" ou simplesmente "reduzir" significará transformar uma matriz por meio de operações elementares com as linhas.)

Algoritmo 3.3 (Eliminação para frente) Dada qualquer matriz A, este algoritmo produz uma forma escalonada de A, inserindo zeros abaixo de cada pivô, trabalhando de cima para baixo.

Passo 1 Descubra a primeira coluna com alguma entrada não nula. Seja j_1 essa coluna.

(a) Arranje as linhas da matriz de tal modo que $a_{1j_1} \neq 0$. Em outras palavras, se necessário, troque a posição de duas linhas para que na primeira linha tenha uma entrada não nula da coluna j_1.

(b) Use a_{1j_1} como um pivô para obter zeros abaixo de a_{1j_1}.
Mais precisamente, para $i > 1$,

(1) Denote $m = -a_{ij_1}/a_{1j_1}$; (2) Substitua R_i por $mR_1 + R_i$

[Ou seja, aplique a operação $-(a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$.]

Passo 2 Repita o Passo 1 com a submatriz formada por todas as linhas da matriz exceto a primeira. Para isso, seja j_2 a primeira coluna do subsistema com uma entrada não nula. Assim, ao final do Passo 2, temos $a_{2j_2} \neq 0$.

Passos 3 a *r* Continue o presente processo até que uma submatriz possua apenas linhas nulas.

Enfatizamos que, ao final do algoritmo, os pivôs serão

$$a_{1j_1},\ a_{2j_2},\ \ldots,\ a_{rj_r}$$

onde r denota o número de linhas não nulas na matriz escalonada final.

OBSERVAÇÃO 1 O número m introduzido no Passo 1(b)(1) é denominado *multiplicador*.

$$m = -\frac{a_{ij_1}}{a_{1j_1}} = -\frac{\text{entrada a ser suprimida}}{\text{pivô}}$$

OBSERVAÇÃO 2 Se todos os escalares originais forem números inteiros, podemos evitar frações se substituirmos a operação do Passo 1(b) pela seguinte.

Substituir R_i por $-a_{ij_1}R_1 + a_{1j_1}R_i$.

Algoritmo 3.4 (Eliminação para trás) Dada qualquer matriz $A = [a_{ij}]$ em forma escalonada com pivôs nas entradas

$$a_{1j_1}, \quad a_{2j_2}, \quad \ldots, \quad a_{rj_r};$$

este algoritmo produz a forma canônica por linhas de A.

Passo 1 (a) (Uso de multiplicação por escalar para igualar o último pivô a 1) Multiplique a última linha não nula R_r por $1/a_{rj_r}$.

(b) (Uso de $a_{rj_r} = 1$ para obter zero acima do pivô) Para $i = r - 1, r - 2, ..., 2, 1,$

$$(1)\ \text{Denote } m = -a_{ij_r}; \qquad (2)\ \text{Substitua } R_i \text{ por } mR_r + R_i$$

(Ou seja, aplique a operação $-a_{ij_r}R_r + R_i \to R_i$.)

Passos 2 a *r* – 1 Repita o Passo 1 para as linhas $R_{r-1}, R_{r-2}, ..., R_2$.

Passo *r* (Uso de multiplicação por escalar para igualar o primeiro pivô a 1) Multiplique R_1 por $1/a_{1j_1}$.

Existe uma forma alternativa do Algoritmo 3.4 que descrevemos usando apenas palavras. A descrição formal desse algoritmo é deixada ao leitor como um problema complementar.

Algoritmo 3.4 alternativo Coloque zeros acima dos pivôs, linha a linha, de baixo para cima (em vez de coluna a coluna da direita para a esquerda).

Esse algoritmo alternativo aplicado a uma matriz aumentada M de um sistema de equações lineares essencialmente consiste em resolver uma a uma as incógnitas pivô de baixo para cima.

OBSERVAÇÃO Enfatizamos que a eliminação gaussiana é um processo de dois estágios, como segue.

ESTÁGIO A (ALGORITMO 3.3) Colocar zeros abaixo de cada pivô, começando com a primeira linha R_1 e seguindo para baixo.

ESTÁGIO B (ALGORITMO 3.4) Colocar zeros acima de cada pivô, começando com a última linha R_r e seguindo para cima.

Existe mais um algoritmo, o de *Gauss-Jordan*, que também reduz uma matriz por linhas à sua forma canônica por linhas. A diferença é que Gauss-Jordan coloca zeros tanto abaixo quanto acima de cada pivô à medida em que começa com a primeira linha R_1 e segue para baixo. Apesar de Gauss-Jordan poder até ser mais simples de enunciar e compreender, é muito menos eficiente do que o algoritmo de eliminação gaussiana de dois estágios.

Exemplo 3.11 Considere a matriz $A = \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$.

(a) Use o Algoritmo 3.3 para reduzir A a uma forma escalonada.
(b) Use o Algoritmo 3.4 para reduzir A até sua forma canônica por linhas.
(a) Primeiro tomamos $a_{11} = 1$ como um pivô para obter zeros abaixo de a_{11}, ou seja, aplicamos as operações "Substituir R_2 por $-2R_1 + R_2$" e "Substituir R_3 por $-3R_1 + R_3$". Depois usamos $a_{23} = 2$ como pivô para obter um zero abaixo de a_{23}, ou seja, aplicamos a operação "Substituir R_3 por $-\frac{3}{2}R_2 + R_3$". Isso fornece

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

Agora, a matriz está em forma escalonada.

(b) Multiplicamos R_3 por $-\frac{1}{2}$ para obter a entrada pivô $a_{35} = 1$ e, então, usamos $a_{35} = 1$ como um pivô para obter zeros acima dele, usando as operações "Substituir R_2 por $-6R_3 + R_2$" e, depois, "Substituir R_1 por $-2R_3 + R_1$". Isso fornece

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Multiplicamos R_2 por $\frac{1}{2}$ para obter a entrada pivô $a_{23} = 1$ e, então, usamos $a_{23} = 1$ como um pivô para obter zeros acima dele, usando a operação "Substituir R_1 por $3R_2 + R_1$". Isso fornece

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 7 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

A última matriz é a forma canônica por linhas de A.

Aplicação a sistemas de equações lineares

Uma maneira de resolver um sistema de equações lineares é utilizar sua matriz aumentada M em vez das equações em si. Mais precisamente, reduzimos M a uma forma escalonada (que nos diz se o sistema possui alguma solução) e, depois, voltamos a reduzir M à sua forma canônica por linhas (que, essencialmente, nos dá a solução do sistema original de equações lineares). A justificativa desse processo decorre dos fatos seguintes.

(1) Qualquer operação elementar com as linhas da matriz aumentada M do sistema equivale a aplicar a operação correspondente ao próprio sistema.

(2) O sistema tem alguma solução se, e só se, a forma escalonada da matriz aumentada M não possui alguma linha da forma $(0, 0, ..., 0, b)$, com $b \neq 0$.

(3) O coeficiente de cada variável básica na forma canônica por linhas da matriz aumentada M (exceto nas linhas nulas) é uma entrada pivô igual a 1 e esse é o único elemento não nulo da respectiva coluna; logo, a forma de variáveis livres da solução do sistema de equações lineares é obtida passando, simplesmente, as variáveis livres para o outro lado da igualdade.

A seguir ilustramos esse processo.

Exemplo 3.12 Resolva os sistemas seguintes.

$$\begin{aligned} x_1 + x_2 - 2x_3 + 4x_4 &= 5 \\ 2x_1 + 2x_2 - 3x_3 + x_4 &= 3 \\ 3x_1 + 3x_2 - 4x_3 - 2x_4 &= 1 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - 2x_3 + 3x_4 &= 4 \\ 2x_1 + 3x_2 + 3x_3 - x_4 &= 3 \\ 5x_1 + 7x_2 + 4x_3 + x_4 &= 5 \end{aligned} \qquad \begin{aligned} x + 2y + z &= 3 \\ 2x + 5y - z &= -4 \\ 3x - 2y - z &= 5 \end{aligned}$$

(a) (b) (c)

(a) Reduzimos a matriz aumentada M desse sistema a uma forma escalonada e, depois, à forma canônica por linhas, como segue.

$$M = \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 2 & 2 & -3 & 1 & 3 \\ 3 & 3 & -4 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Reescrevemos a forma canônica por linhas em termos de um sistema de equações lineares para obter a forma de variáveis livres da solução. Isto é,

$$\begin{aligned} x_1 + x_2 \quad - 10x_4 &= -9 \\ x_3 - 7x_4 &= -7 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x_1 &= -9 - x_2 + 10x_4 \\ x_3 &= -7 + 7x_4 \end{aligned}$$

(Omitimos a linha nula na solução.) Observe que x_1 e x_3 são as variáveis pivô e x_2 e x_4 são as variáveis livres.

(b) Primeiro reduzimos a matriz aumentada M a uma forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 2 & 3 & 3 & -1 & 3 \\ 5 & 7 & 4 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 2 & 14 & -14 & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

Não precisamos continuar até obter a forma canônica por linhas de M, porque a forma escalonada já nos diz que o sistema não possui solução. De fato, a terceira linha da matriz escalonada corresponde à equação degenerada

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -5$$

que não possui solução. Assim, o sistema não possui solução.

(c) Reduzimos a matriz aumentada M a uma forma escalonada e depois à forma canônica por linhas, como segue.

$$M = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & -1 & -4 \\ 3 & -2 & -1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & -8 & -4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & -28 & -84 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Assim, o sistema tem a única solução $x = 2$, $y = -1$, $z = 3$ ou, equivalentemente, o vetor $u = (2, -1, 3)$. Observe que a forma escalonada de M já indica que a solução é única, pois corresponde a um sistema triangular.

Aplicação a teoremas de existência e unicidade

Nesta subseção discutimos as condições teóricas para a existência e unicidade de uma solução de um sistema de equações lineares usando o conceito de posto de uma matriz.

Teorema 3.9 Considere um sistema de equações lineares com n incógnitas e matriz aumentada $M = [A, B]$. Então

(a) o sistema possui alguma solução se, e só se, $\text{pos}(A) = \text{pos}(M)$,
(b) a solução é única se, e só se, $\text{pos}(A) = \text{pos}(M) = n$.

Demonstração de (a) O sistema tem alguma solução se, e só se, alguma forma escalonada de $M = [A, B]$ não possui uma linha no formato

$$(0, 0, ..., 0, b), \text{ com } b \neq 0$$

Se alguma forma escalonada de M possui uma tal linha, então b é um pivô de M, mas não de A e, portanto, $\text{pos}(M) > \text{pos}(A)$. Caso contrário, as formas escalonadas de A e de M têm os mesmos pivôs e, portanto, $\text{pos}(A) = \text{pos}(M)$, provando (a).

Demonstração de (b): O sistema possui uma única solução se, e só se, alguma forma escalonada não possui variáveis livres. Isso significa que existe um pivô para cada incógnita. Assim, $n = \text{pos}(A) = \text{pos}(M)$, provando (b).

Nessa demonstração utilizamos o fato (Problema 3.74) que uma forma escalonada de uma matriz aumentada $M = [A, B]$ automaticamente fornece uma forma escalonada de A.

3.9 EQUAÇÃO MATRICIAL DE UM SISTEMA DE EQUAÇÕES LINEARES

O sistema geral (3.2) de m equações lineares com n incógnitas é equivalente à equação matricial

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \quad \text{ou} \quad AX = B$$

onde $A = [a_{ij}]$ é a matriz de coeficientes, $X = [x_j]$ é o vetor coluna das incógnitas e $B = [b_i]$ é o vetor coluna dos termos constantes. (Em alguns livros pode ser encontrada a equação $Ax = b$ em vez de $AX = B$, para enfatizar que x e b são, simplesmente, vetores coluna.)

Afirmar que o sistema de equações lineares e a equação matricial são equivalentes significa que qualquer vetor solução do sistema é uma solução da equação matricial e vice-versa.

Exemplo 3.13 São equivalentes o sistema de equações lineares e a equação matricial seguintes.

$$\begin{aligned} x_1 + 2x_2 - 4x_3 + 7x_4 &= 4 \\ 3x_1 - 5x_2 + 6x_3 - 8x_4 &= 8 \\ 4x_1 - 3x_2 - 2x_3 + 6x_4 &= 11 \end{aligned} \quad \text{e} \quad \begin{bmatrix} 1 & 2 & -4 & 7 \\ 3 & -5 & 6 & -8 \\ 4 & -3 & -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 11 \end{bmatrix}$$

Observe que $x_1 = 3$, $x_2 = 1$, $x_3 = 2$, $x_4 = 1$ ou, equivalentemente, o vetor $u = [3, 1, 2, 1]$ é uma solução do sistema. Assim, o vetor (coluna) u também é uma solução da equação matricial.

A forma matricial $AX = B$ de um sistema de equações lineares é muito conveniente em termos de escrita quando discutimos e demonstramos propriedades de sistemas de equações lineares. Exemplificamos isso com nosso primeiro teorema (descrito na Figura 3-1), que reformulamos como segue.

Teorema 3.1 Seja K um corpo infinito. Então o sistema $AX = B$ tem (a) uma única solução, (b) nenhuma solução, ou (c) uma infinidade de soluções.

Demonstração Basta demonstrar que, se $AX = B$ tiver mais do que uma solução, então terá uma infinidade de soluções. Suponha que u e v sejam soluções distintas de $AX = B$, ou seja, que $Au = B$ e $Av = B$. Então, dado qualquer r de K,

$$A[u + r(u - v)] = Au + r(Au - Av) = B + r(B - B) = B$$

Assim, para cada r de K, o vetor $u + r(u - v)$ é uma solução de $AX = B$. Como todas essas soluções são distintas (Problema 3.47), $AX = B$ possui uma infinidade de soluções.

Observe que esse teorema é verdadeiro se K for o corpo dos números reais **R** (ou o corpo complexo **C**). Na Seção 3.3 mostramos que esse teorema tem uma interpretação geométrica se o sistema consistir em duas equações com duas incógnitas, em que cada equação representa uma reta em $\mathbf{R}^2$. O teorema também tem uma interpretação geométrica se o sistema consistir em três equações não degeneradas com três incógnitas, em que as três equações representam planos H_1, H_2 e H_3 em $\mathbf{R}^3$, como segue.

(a) *Solução única* Os três planos se cortam exatamente num ponto.
(b) *Nenhuma solução* Os planos podem se intersectar dois a dois, mas não possuir ponto comum de interseção, ou dois dos planos podem ser paralelos e distintos.
(c) *Uma infinidade de soluções* Os três planos podem se intersectar numa reta (uma variável livre) ou coincidem (duas variáveis livres).

Os três casos estão ilustrados na Figura 3-3.

Equação matricial de um sistema de equações lineares quadrado

Se a matriz A de coeficientes for quadrada, dizemos que o sistema de equações lineares $AX = B$ é quadrado. Nesse caso, temos o resultado importante seguinte.

(*a*) Solução única

(*c*) Uma infinidade de soluções

(*b*) Nenhuma solução

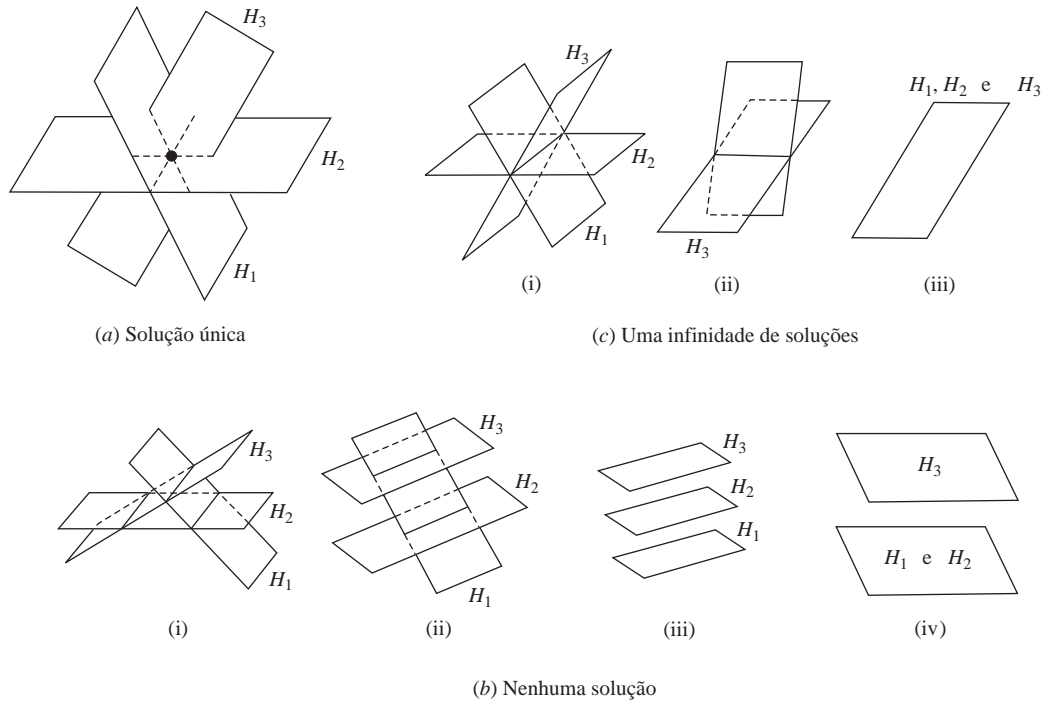



Figura 3-3

Teorema 3.10 Um sistema quadrado $AX = B$ de equações lineares tem uma única solução se, e só se, a matriz A é invertível. Nesse caso, $A^{-1}B$ é a única solução do sistema.

Aqui apenas demonstramos que se A for invertível, então $A^{-1}B$ é a única solução. Se A for invertível, então

$$A(A^{-1}B) = (AA^{-1})B = IB = B$$

e, portanto, $A^{-1}B$ é uma solução. Seja, agora, v uma solução qualquer, ou seja, $Av = B$. Então

$$v = Iv = (A^{-1}A)v = A^{-1}(Av) = A^{-1}B$$

Assim, a solução $A^{-1}B$ é única.

Exemplo 3.14 Considere o seguinte sistema de equações lineares, com matriz de coeficientes A e inversa A^{-1} também dadas.

$$\begin{aligned} x + 2y + \ \ 3z &= 1 \\ x + 3y + \ \ 6z &= 3 \\ 2x + 6y + 13z &= 5 \end{aligned}, \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 6 \\ 2 & 6 & 13 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix}$$

Pelo Teorema 3.10, a única solução do sistema é

$$A^{-1}B = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \\ -1 \end{bmatrix}$$

Isto é, $x = -6$, $y = 5$, $z = -1$.

OBSERVAÇÃO Enfatizamos que, em geral, o Teorema 3.10 não nos ajuda a encontrar a solução de um sistema quadrado. Isso ocorre porque, em geral, encontrar a inversa de uma matriz de coeficientes A não é mais fácil do que resolver o sistema diretamente. Assim, a menos que nos seja dada a inversa de uma matriz de coeficientes A, como no exemplo precedente, resolvemos um sistema quadrado usando a eliminação gaussiana (ou algum método iterativo cuja discussão está fora dos objetivos deste texto.)

3.10 SISTEMAS DE EQUAÇÕES LINEARES E COMBINAÇÃO LINEAR DE VETORES

O sistema geral (3.2) de equações lineares pode ser reescrito como a equação vetorial seguinte.

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{bmatrix}$$

Lembre que um vetor v de K^n é dito uma *combinação linear* dos vetores $u_1, u_2, ..., u_m$ de K^n se existirem escalares $a_1, a_2, ..., a_m$ em K tais que

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

Em vista disso, o sistema geral (3.2) de equações lineares e a equação vetorial equivalente dada têm alguma solução se, e só se, o vetor coluna dos termos constantes é uma combinação linear das colunas da matriz de coeficientes. Enunciamos essa observação formalmente.

Teorema 3.11 Um sistema $AX = B$ de equações lineares possui alguma solução se, e só se, B é uma combinação linear das colunas da matriz de coeficientes A.

Assim, o problema de expressar um dado vetor v de K^n como uma combinação linear de vetores $u_1, u_2, ..., u_m$ de K^n se reduz à resolução de um sistema de equações lineares.

Exemplo de combinação linear

Digamos que queiramos expressar o vetor $v = (1, -2, 5)$ como uma combinação linear dos vetores

$$u_1 = (1, 1, 1), \qquad u_2 = (1, 2, 3), \qquad u_3 = (2, -1, 1)$$

Começamos escrevendo $v = xu_1 + yu_2 + zu_3$ com incógnitas x, y, z e depois encontramos o sistema de equações lineares equivalente que resolvemos. Mais precisamente, primeiro escrevemos

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + z \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \qquad (*)$$

Então

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = \begin{bmatrix} x \\ x \\ x \end{bmatrix} + \begin{bmatrix} y \\ 2y \\ 3y \end{bmatrix} + \begin{bmatrix} 2z \\ -z \\ z \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ x + 2y - z \\ x + 3y + z \end{bmatrix} \qquad (**)$$

Igualando entradas correspondentes, obtemos o sistema equivalente

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned}$$

Por conveniência notacional, escrevemos os vetores de $\mathbf{R}^n$ como colunas, porque dessa forma é mais fácil encontrar o sistema de equações lineares equivalente. De fato, é fácil ir diretamente da equação vetorial (*) para o sistema (**).

Agora resolvemos o sistema de equações lineares equivalente por redução à forma escalonada. Isso fornece

$$\begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \qquad \text{e, então,} \qquad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

Uma substituição para trás fornece a solução $x = -6, y = 3, z = 2$. Assim, $v = -6u_1 + 3u_2 + 2u_3$.

Exemplo 3.15

(a) Escreva o vetor $v = (4, 9, 19)$ como uma combinação linear de

$$u_1 = (1, -2, 3), \qquad u_2 = (3, -7, 10), \qquad u_3 = (2, 1, 9).$$

Encontramos o sistema de equações lineares equivalente escrevendo $v = xu_1 + yu_2 + zu_3$ e reduzimos o sistema por redução à forma escalonada. Temos

$$\begin{aligned} x + 3y + 2z &= 4 \\ -2x - 7y + z &= 9 \\ 3x + 10y + 9z &= 19 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 3y + 2z &= 4 \\ -y + 5z &= 17 \\ y + 3z &= 7 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 3y + 2z &= 4 \\ -y + 5z &= 17 \\ 8z &= 24 \end{aligned}$$

A substituição para trás fornece a solução $x = 4$, $y = -2$, $z = 3$. Assim, v é uma combinação linear de u_1, u_2, u_3. Mais precisamente, $v = 4u_1 - 2u_2 + 3u_3$.

(b) Escreva o vetor $v = (2, 3, -5)$ como uma combinação linear de

$$u_1 = (1, 2, -3), \qquad u_2 = (2, 3, -4), \qquad u_3 = (1, 3, -5)$$

Encontramos o sistema de equações lineares equivalente escrevendo $v = xu_1 + yu_2 + zu_3$ e reduzimos o sistema por redução à forma escalonada. Temos

$$\begin{aligned} x + 2y + z &= 2 \\ 2x + 3y + 3z &= 3 \\ -3x - 4y - 5z &= -5 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ -y + z &= -1 \\ 2y - 2z &= 1 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ -5y + 5z &= -1 \\ 0 &= 3 \end{aligned}$$

Esse sistema não tem solução. Assim, é impossível escrever v como uma combinação linear de u_1, u_2, u_3.

Combinações lineares de vetores ortogonais, coeficientes de Fourier

Inicialmente revemos (Seção 1.4) que o produto escalar $u \cdot v$ dos vetores $u = (a_1, a_2, ..., a_n)$ e $v = (b_1, b_2, ..., b_n)$ de $\mathbf{R}^n$ é definido por

$$u \cdot v = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

Além disso, os dois vetores u e v são ditos *ortogonais* se seu produto escalar for nulo, $u \cdot v = 0$.

Suponha que $u_1, u_2, ..., u_n$ de $\mathbf{R}^n$ sejam n vetores não nulos dois a dois ortogonais. Isso significa que

$$\text{(i)} \quad u_i \cdot u_j = 0 \quad \text{com } i \neq j \qquad \text{e} \qquad \text{(ii)} \quad u_i \cdot u_i \neq 0 \quad \text{para cada } i$$

Então, para qualquer vetor v de $\mathbf{R}^n$, existe uma maneira simples de escrever v como combinação linear de $u_1, u_2, ..., u_n$, que apresentamos no exemplo seguinte.

Exemplo 3.16 Considere os três vetores de $\mathbf{R}^3$.

$$u_1 = (1, 1, 1), \qquad u_2 = (1, -3, 2), \qquad u_3 = (5, -1, -4)$$

Esses vetores são dois a dois ortogonais, ou seja,

$$u_1 \cdot u_2 = 1 - 3 + 2 = 0, \qquad u_1 \cdot u_3 = 5 - 1 - 4 = 0, \qquad u_2 \cdot u_3 = 5 + 3 - 8 = 0$$

Suponha que queiramos escrever $v = (4, 14, -9)$ como combinação linear de u_1, u_2, u_3.

Método 1 Encontre o sistema de equações lineares equivalente, como no Exemplo 3.14 e resolva-o, encontrando $v = 3u_1 - 4u_2 + u_3$.

Método 2 (Neste método usamos a ortogonalidade dos vetores u_1, u_2, u_3 e, por isso, é muito mais simples.) Escreva v como combinação linear de u_1, u_2, u_3 usando escalares x, y, z a determinar, como segue.

$$(4, 14, -9) = x(1, 1, 1) + y(1, -3, 2) + z(5, -1, -4) \qquad (*)$$

Tomando o produto escalar de (*) com u_1, obtemos

$$(4, 14, -9) \cdot (1, 1, 1) = x(1, 1, 1) \cdot (1, 1, 1) \quad \text{ou} \quad 9 = 3x \quad \text{ou} \quad x = 3$$

(As duas últimas parcelas do produto escalar desaparecem porque u_1 é ortogonal a u_2 e a u_3.) Tomando, em seguida, o produto escalar de (*) com u_2, obtemos

$$(4, 14, -9) \cdot (1, -3, 2) = y(1, -3, 2) \cdot (1, -3, 2) \quad \text{ou} \quad -56 = 14y \quad \text{ou} \quad y = -4$$

Finalmente, tomando o produto escalar de (*) com u_3, obtemos

$$(4, 14, -9) \cdot (5, -1, -4) = z(5, -1, -4) \cdot (5, -1, -4) \quad \text{ou} \quad 42 = 42z \quad \text{ou} \quad z = 1$$

Assim, $v = 3u_1 - 4u_2 + u_3$.

O procedimento do Método 2 é válido em geral, como segue.

Teorema 3.12 Sejam $u_1, u_2, ..., u_n$ vetores não nulos dois a dois ortogonais de $\mathbf{R}^n$. Então, para qualquer vetor v de $\mathbf{R}^n$,

$$v = \frac{v \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{v \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{v \cdot u_n}{u_n \cdot u_n} u_n$$

Enfatizamos que precisamos de n tais vetores ortogonais u_i de $\mathbf{R}^n$ para poder usar a fórmula. Observe, também, que cada $u_i \cdot u_i \neq 0$, por ser u_i um vetor não nulo.

OBSERVAÇÃO Dizemos que os escalares que aparecem no Teorema 3.12 são os *coeficientes de Fourier* de v. Mais precisamente, o coeficiente de Fourier de v em relação a u_i é

$$k_i = \frac{v \cdot u_i}{u_i \cdot u_i} = \frac{v \cdot u_i}{\|u_i\|^2}$$

Esses coeficientes são análogos aos famosos coeficientes da série de Fourier de uma função.

3.11 SISTEMAS HOMOGÊNEOS DE EQUAÇÕES LINEARES

Um sistema de equações lineares é dito *homogêneo* se todos os seus termos constantes forem nulos. Assim, um sistema homogêneo tem a forma $AX = 0$. Claramente, tais sistemas sempre têm o vetor nulo $0 = (0, 0, ..., 0)$ como uma solução, denominada *solução nula* ou *trivial*. Por isso, geralmente estamos interessados em saber se esses sistemas têm, ou não, alguma solução não nula.

Como o sistema homogêneo $AX = 0$ tem, pelo menos, a solução nula, sempre pode ser colocado em forma escalonada, digamos,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= 0 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= 0 \end{aligned}$$

Aqui, r denota o número de equações na forma escalonada e n denota o número de incógnitas. Assim, o sistema escalonado tem $n - r$ variáveis livres.

A questão de soluções não nulas se reduz aos dois casos seguintes.

(i) $r = n$. O sistema só tem a solução nula.
(ii) $r < n$. O sistema tem alguma solução não nula.

Por isso, começando com menos equações do que incógnitas, temos $r < n$ na forma escalonada e o sistema terá alguma solução não nula. Isso demonstra o resultado importante a seguir.

Teorema 3.13 Um sistema homogêneo $AX = 0$ com mais incógnitas do que equações tem alguma solução não nula.

Exemplo 3.17 Decida se os sistemas homogêneos dados têm ou não solução não nula.

$$\begin{array}{c}\begin{aligned} x + y - z &= 0\\ 2x - 3y + z &= 0\\ x - 4y + 2z &= 0 \end{aligned}\\ \text{(a)}\end{array}\qquad \begin{array}{c}\begin{aligned} x + y - z &= 0\\ 2x + 4y - z &= 0\\ 3x + 2y + 2z &= 0 \end{aligned}\\ \text{(b)}\end{array}\qquad \begin{array}{c}\begin{aligned} x_1 + 2x_2 - 3x_3 + 4x_4 &= 0\\ 2x_1 - 3x_2 + 5x_3 - 7x_4 &= 0\\ 5x_1 + 6x_2 - 9x_3 + 8x_4 &= 0 \end{aligned}\\ \text{(c)}\end{array}$$

(a) Reduzimos o sistema à forma escalonada, como segue.

$$\begin{aligned} x + y - z &= 0\\ -5y + 3z &= 0\\ -5y + 3z &= 0 \end{aligned} \qquad \text{e, então,} \qquad \begin{aligned} x + y - z &= 0\\ -5y + 3z &= 0 \end{aligned}$$

Esse sistema tem uma solução não nula porque a forma escalonada tem somente duas equações nas três variáveis. Aqui, z é uma variável livre. Tomando, digamos, $z = 5$, a substituição para trás dá $y = 3$ e $x = 2$. Assim, o vetor $u = (2, 3, 5)$ é uma solução não nula específica.

(b) Reduzimos o sistema à forma escalonada, como segue.

$$\begin{aligned} x + y - z &= 0\\ 2y + z &= 0\\ -y + 5z &= 0 \end{aligned} \qquad \text{e, então,} \qquad \begin{aligned} x + y - z &= 0\\ 2y + z &= 0\\ 11z &= 0 \end{aligned}$$

A forma escalonada tem três equações com três variáveis. Assim, o sistema só tem a solução nula.

(c) O sistema precisa ter alguma solução não nula (Teorema 3.13) porque há quatro incógnitas, mas só três equações. (Aqui não precisamos reduzir o sistema a uma forma escalonada.)

Base para o conjunto solução de um sistema homogêneo

Seja W o conjunto de todas as soluções de um sistema homogêneo $AX = 0$. Uma lista $u_1, u_2, ..., u_s$ de vetores solução não nulos do sistema é uma *base de W* se cada vetor solução $w \in W$ pode ser expresso, de maneira única, como uma combinação linear dos vetores $u_1, u_2, ..., u_s$, ou seja, se existem escalares únicos $a_1, a_2, ..., a_s$ tais que

$$w = a_1u_1 + a_2u_2 + \cdots + a_su_s$$

O número s de tais vetores da base é igual ao número de variáveis livres. Este número s é dito a *dimensão* de W, que denotamos por $\dim W = s$. Caso $W = \{0\}$, ou seja, se o sistema possuir apenas a solução nula, definimos $\dim W = 0$.

O teorema seguinte, demonstrado no Capítulo 5, nos diz como determinar uma tal base.

Teorema 3.14 Seja W o conjunto das soluções de um sistema homogêneo $AX = 0$ e suponha que a forma escalonada do sistema homogêneo tenha s variáveis livres. Sejam $u_1, u_2, ..., u_s$ as soluções obtidas tomando uma das variáveis livres como 1 (ou qualquer outra constante não nula) e as demais variáveis livres como 0. Então $\dim W = s$ e $u_1, u_2, ..., u_s$ formam uma base de W.

Enfatizamos que o conjunto das soluções W pode ter muitas bases e que o Teorema 3.14 apenas dá uma delas.

Exemplo 3.18 Encontre a dimensão e uma base do conjunto solução W do sistema homogêneo

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0\\ 2x_1 + 4x_2 - 5x_3 + x_4 - 6x_5 &= 0\\ 5x_1 + 10x_2 - 13x_3 + 4x_4 - 16x_5 &= 0 \end{aligned}$$

Primeiro reduza o sistema a uma forma escalonada, aplicando as operações seguintes.

"Substituir L_2 por $-2L_1 + L_2$", "Substituir L_3 por $-5L_1 + L_3$" e, depois, "Substituir L_3 por $-2L_2 + L_3$"

Essas operações fornecem

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \\ 2x_3 - 6x_4 + 4x_5 &= 0 \end{aligned} \qquad \text{e} \qquad \begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \end{aligned}$$

O sistema em forma escalonada tem três variáveis livres, x_2, x_4 e x_5, portanto, dim $W = 3$. Três vetores solução que formam uma base de W são obtidos como segue.

(1) Tomamos $x_2 = 1$, $x_4 = 0$ e $x_5 = 0$. Substituição para trás fornece a solução $v_1 = (-2, 1, 0, 0, 0)$.
(2) Tomamos $x_2 = 0$, $x_4 = 1$ e $x_5 = 0$. Substituição para trás fornece a solução $v_2 = (7, 0, 3, 1, 0)$.
(3) Tomamos $x_2 = 0$, $x_4 = 0$ e $x_5 = 1$. Substituição para trás fornece a solução $v_3 = (-2, 0, -2, 0, 1)$.

Os vetores solução $v_1 = (-2, 1, 0, 0, 0)$, $v_2 = (7, 0, 3, 1, 0)$, $v_3 = (-2, 0, -2, 0, 1)$ formam uma base de W.

OBSERVAÇÃO Qualquer solução do sistema do Exemplo 3.18 pode ser escrita da forma

$$\begin{aligned} av_1 + bv_2 + cv_3 &= a(-2, 1, 0, 0, 0) + b(7, 0, 3, 1, 0) + c(-2, 0, -2, 0, 1) \\ &= (-2a + 7b - 2c, \quad a, \quad 3b - 2c, \quad b, \quad c) \end{aligned}$$

ou

$$x_1 = -2a + 7b - 2c, \qquad x_2 = a, \qquad x_3 = 3b - 2c, \qquad x_4 = b, \qquad x_5 = c$$

em que a, b, c são constantes arbitrárias. Observe que essa representação não é nada mais do que a forma paramétrica da solução geral, com a escolha dos parâmetros $x_2 = a$, $x_4 = b$, $x_5 = c$.

Sistemas não homogêneos e sistemas homogêneos associados

Seja $AX = B$ um sistema não homogêneo de equações lineares. Dizemos que $AX = 0$ é o *sistema homogêneo associado*. Por exemplo,

$$\begin{aligned} x + 2y - 4z &= 7 \\ 3x - 5y + 6z &= 8 \end{aligned} \qquad \text{e} \qquad \begin{aligned} x + 2y - 4z &= 0 \\ 3x - 5y + 6z &= 0 \end{aligned}$$

são um sistema não homogêneo e o sistema homogêneo associado.

A relação entre uma solução U de um sistema não homogêneo $AX = B$ e o conjunto solução W do sistema homogêneo associado $AX = 0$ é dada no teorema seguinte.

Teorema 3.15 Sejam v_0 uma solução particular de $AX = B$ e W o conjunto solução de $AX = 0$. Então o conjunto solução de $AX = B$ é dado por

$$U = v_0 + W = \{v_0 + w : w \in W\}$$

Em outras palavras, $U = v_0 + W$ é obtido somando v_0 a cada elemento de W. Observamos que esse teorema tem uma interpretação geométrica em $\mathbf{R}^3$, como segue. Suponha que W seja uma reta passando pela origem. Então, conforme esboçado na Figura 3-4, $U = v_0 + W$ é a reta paralela a W obtida pela soma de v_0 a cada elemento de W. Analogamente, se W for um plano pela origem O, então $U = v_0 + W$ é um plano paralelo a W.

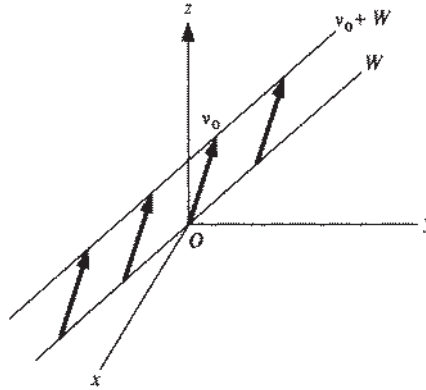



Figura 3-4

3.12 MATRIZES ELEMENTARES

Sejam e uma operação elementar com as linhas e $e(A)$ o resultado dessa operação a uma matriz A. Seja E a matriz obtida pela aplicação de e à matriz identidade I, ou seja,

$$E = e(I)$$

Então E é denominada *matriz elementar* correspondente à operação elementar e. Observe que E sempre é uma matriz quadrada.

Exemplo 3.19 Considere as operações elementares com as linhas a seguir.

(1) Trocar R_2 e R_3 entre si. (2) Substituir R_2 por $-6R_2$. (3) Substituir R_3 por $-4R_1 + R_3$.

As matrizes elementares 3×3 correspondentes a essas operações elementares com as linhas são

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$$

Temos o teorema seguinte, demonstrado no Problema 3.34.

Teorema 3.16 Sejam e uma operação elementar com as linhas e E a matriz elementar $m \times m$ correspondente. Então

$$e(A) = EA$$

onde A é uma matriz $m \times n$ qualquer.

Em outras palavras, o resultado de aplicar uma operação elementar e nas linhas de A pode ser obtido multiplicando a matriz elementar correspondente E pela matriz A, nessa ordem.

Agora suponha que e' seja a operação inversa da operação elementar e e que E' e E sejam as matrizes correspondentes. Observamos (Problema 3.33) que E é invertível e que E' é sua inversa. Isso significa, em particular, que qualquer produto

$$P = E_k \ldots E_2 E_1$$

de matrizes elementares é invertível.

Aplicações de matrizes elementares

Usando o Teorema 3.16, estamos em condições de provar (Problema 3.35) as seguintes propriedades importantes de matrizes.

Teorema 3.17 Seja A uma matriz quadrada. Então as afirmativas seguintes são equivalentes.

(a) A é invertível (não singular).
(b) A é equivalente por linhas à matriz identidade I.
(c) A é um produto de matrizes elementares.

Lembre que duas matrizes A e B são ditas inversas se $AB = BA = I$. O teorema seguinte (demonstrado no Problema 3.36) demonstra que só precisamos verificar que um desses produtos é igual à identidade, digamos, $AB = I$, para estabelecer que as matrizes são inversas.

Teorema 3.18 Suponha que $AB = I$. Então $BA = I$ e, portanto, $B = A^{-1}$.

A equivalência por linhas também pode ser definida em termos da multiplicação matricial. Mais precisamente, demonstra-se (Problema 3.37) o teorema seguinte.

Teorema 3.19 B é equivalente por linhas a A se, e só se, existe uma matriz não singular P tal que $B = PA$.

Aplicação à obtenção da inversa de uma matriz $n \times n$

A inversa de uma matriz pode ser encontrada com o algoritmo seguinte.

Algoritmo 3.5 Dada qualquer matriz A, este algoritmo produz a inversa de A, ou mostra que a inversa não existe.

Passo 1 Construa a matriz $n \times 2n$ (em blocos) $M = [A, I]$, em que A é a metade da esquerda e a matriz identidade I é a metade da direita de M.

Passo 2 Reduza M por linhas a uma forma escalonada. Se nesse processo surgir uma linha nula na metade A de M, então

PARE

A não possui inversa. (Caso contrário, A está em forma triangular.)

Passo 3 Continue reduzindo M à sua forma canônica por linhas

$$M \sim [I, B]$$

onde a matriz identidade I substituiu A na metade da esquerda de M.

Passo 4 Defina $A^{-1} = B$, a matriz que agora está na metade da direita de M.

A justificativa para esse processo é a seguinte. Suponha que A seja invertível e que, digamos, a metade esquerda de M, que é A, tenha sido reduzida à matriz identidade I com a sequência $e_1, e_2, ..., e_q$ de operações elementares com as linhas de $M = [A, I]$. Seja E_i a matriz elementar correspondente à operação elementar e_i. Então, pelo Teorema 3.16, obtemos

$$E_q \dots E_2 E_1 A = I \qquad \text{ou} \qquad (E_q \dots E_2 E_1 I)A = I, \quad \text{portanto,} \quad A^{-1} = E_q \dots E_2 E_1 I$$

Isso mostra que A^{-1} pode ser obtida pela aplicação das operações elementares $e_1, e_2, ..., e_q$ à matriz identidade I, que colocamos na metade da direita de M. Assim, $B = A^{-1}$, como afirmamos.

Exemplo 3.20 Obtenha a inversa da matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$.

Primeiro formamos a matriz (em blocos) $M = [A, I]$ e a reduzimos a uma forma escalonada.

$$M = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right]$$

Na forma escalonada, a metade da esquerda de M está em forma triangular, portanto, A tem inversa. Em seguida, continuamos reduzindo M à forma canônica por linhas.

$$M \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

Agora, a matriz identidade está na metade da esquerda da matriz final, de modo que a metade da direita é A^{-1}. Em outras palavras,

$$A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$$

Operações elementares com as colunas

Consideremos, agora, uma matriz A com colunas $C_1, C_2, ..., C_n$. As seguintes operações sobre A, análogas às operações elementares com as linhas, são denominadas *operações elementares com as colunas*.

[F_1] (Troca da Ordem de Colunas) Trocar as colunas C_i e C_j entre si.
[F_2] (Multiplicação de Coluna por Escalar) Substituir a coluna C_i por kC_i (com $k \neq 0$).
[F_3] (Soma de Colunas) Substituir a coluna C_j por $kC_i + C_j$.

Podemos indicar cada uma dessas operações com as colunas escrevendo, respectivamente,

$$(1)\ C_i \leftrightarrow C_j, \qquad (2)\ kC_i \to C_i, \qquad (3)\ (kC_i + C_j) \to C_j$$

Além disso, cada operação com as colunas tem uma operação inversa do mesmo tipo, exatamente como ocorre com as operações com as linhas.

Sejam, agora, f uma operação elementar com as colunas e F a matriz obtida pela aplicação de f à matriz identidade I, ou seja,

$$F = f(I)$$

Então F é denominada *matriz elementar* correspondente à operação elementar f. Observe que F sempre é uma matriz quadrada.

Exemplo 3.21 Considere as operações elementares com as colunas a seguir.

(1) Trocar C_1 e C_3 entre si. (2) Substituir C_3 por $-2C_3$. (3) Substituir C_3 por $-3C_2 + C_3$.

As matrizes elementares 3×3 correspondentes a essas operações elementares com as colunas são

$$F_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad F_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \qquad F_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

O teorema seguinte é análogo ao Teorema 3.16 relativo às operações elementares com as linhas.

Teorema 3.20 Dada qualquer matriz A, temos $f(A) = AF$.

Em outras palavras, o resultado de aplicar uma operação elementar f nas colunas de A pode ser obtido multiplicando a matriz A pela matriz elementar correspondente F, nessa ordem.

Equivalência de matrizes

Uma matriz B é equivalente a uma matriz A se B puder ser obtida a partir de A por uma sequência de operações com as linhas e as colunas. Equivalentemente, B é equivalente a A se existirem matrizes não singulares P e Q tais que $B = PAQ$. Da mesma forma que a equivalência por linhas, a equivalência de matrizes é uma relação de equivalência.

O resultado principal desta subseção (demonstrado no Problema 3.38) é o seguinte.

Teorema 3.21 Cada matriz $m \times n$ A é equivalente a uma única matriz em blocos da forma

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

em que I_r é a matriz identidade de ordem r.

Temos a definição seguinte.

DEFINIÇÃO O inteiro não negativo r no Teorema 3.21 é o *posto* de A, denotado por pos(A).

Observe que essa definição concorda com a definição já dada de posto de uma matriz.

3.13 DECOMPOSIÇÃO *LU*

Seja A uma matriz não singular que pode ser levada à forma triangular (superior) U usando somente operações de soma com as linhas, ou seja, A pode ser triangulada pelo algoritmo seguinte, que escrevemos na notação computacional.

Algoritmo 3.6 Dada qualquer matriz A, este algoritmo produz uma matriz triangular U.

Passo 1 Repita para $i = 1, 2, ..., n - 1$

Passo 2 Repita para $j = i + 1, i + 2, ..., n$

(a) Tome $m_{ij} := -a_{ij}/a_{ii}$.

(b) Tome $R_j := m_{ij}R_i + R_j$

[Fim do Passo 2, circuito interno.]

[Fim do Passo 1, circuito externo.]

Dizemos que os números m_{ij} são *multiplicadores*. Às vezes, mantemos esses multiplicadores controlados pela matriz triangular inferior L seguinte.

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -m_{21} & 1 & 0 & \dots & 0 & 0 \\ -m_{31} & -m_{32} & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -m_{n1} & -m_{n2} & -m_{n3} & \dots & -m_{n,n-1} & 1 \end{bmatrix}$$

Assim, L possui entradas 1 na diagonal, zeros acima da diagonal e o simétrico do multiplicador m_{ij} como sua ij-entrada abaixo da diagonal.

Essa matriz L e a matriz triangular U obtida no Algoritmo 3.6 nos fornecem a clássica fatoração LU da matriz A. Temos o resultado seguinte.

Teorema 3.22 Seja A uma matriz não singular que pode ser levada à forma triangular usando somente operações de soma com as linhas. Então $A = LU$, onde L é a matriz triangular inferior com entradas 1 na diagonal precedente e U é uma matriz triangular superior com zeros na diagonal.

Exemplo 3.22 Seja $A = \begin{bmatrix} 1 & 2 & -3 \\ -3 & -4 & 13 \\ 2 & 1 & -5 \end{bmatrix}$. Observe que A pode ser reduzida à forma triangular com as operações "Substituir R_2 por $3R_1 + R_2$", "Substituir R_3 por $-2R_1 + R_3$" e, por último, "Substituir R_3 por $\frac{3}{2}R_2 + R_3$", como segue.

$$A \sim \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{bmatrix}$$

Isso dá a fatoração $A = LU$ clássica, com

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{2} & 1 \end{bmatrix} \qquad \text{e} \qquad U = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{bmatrix}$$

Destacamos que

(1) as entradas -3, 2 e $-\frac{3}{2}$ de L são os simétricos dos multiplicadores das operações elementares com as linhas dadas e

(2) U é a forma triangular de A.

Aplicações a sistemas de equações lineares

Seja M um algoritmo computacional e denotemos por $C(n)$ seu tempo de processamento como uma função do tamanho n de seus dados iniciais. [A função $C(n)$ costuma ser denominada *complexidade temporal* do algoritmo M ou, simplesmente, sua *complexidade*.] Frequentemente, $C(n)$ simplesmente conta o número de produtos e quocientes executados por M, mas não conta o número de somas e subtrações, porque essas levam muito menos tempo.

Seja, agora, $AX = B$ um sistema quadrado de equações lineares, com

$$A = [a_{ij}], \qquad X = [x_1, \dots, x_n]^T, \qquad B = [b_1, \dots, b_n]^T$$

e suponha que A tenha uma fatoração LU. Então o sistema pode ser levado a uma forma triangular pelo Algoritmo 3.6 aplicado à matriz aumentada $M = [A, B]$ do sistema (e depois efetuar a substituição para trás). A complexidade temporal do Algoritmo 3.6 e da substituição para trás são, respectivamente,

$$C(n) \approx \tfrac{1}{2}n^3 \qquad \text{e} \qquad C(n) \approx \tfrac{1}{2}n^2$$

onde n é a quantidade de equações.

Por outro lado, suponha que já tenhamos a fatoração $A = LU$. Então, para triangular o sistema, basta aplicar as operações com as linhas do algoritmo (cuja memória está em L) ao vetor coluna B. Nesse caso, a complexidade temporal é

$$C(n) \approx \tfrac{1}{2}n^2$$

É claro que, para obter a fatoração $A = LU$, precisamos do algoritmo original, em que $C(n) \approx \frac{1}{2}n^3$. Assim, possivelmente nada é ganho calculando primeiro a fatoração LU no caso de um único sistema. Contudo, há situações, como a apresentada a seguir, em que a fatoração LU tem utilidade.

Suponha que, para uma dada matriz A dada, necessitemos resolver o sistema

$$AX = B$$

repetidamente, para uma sequência de diferentes vetores constantes, digamos, $B_1, B_2, \dots, B_k$. Suponha, também, que alguns desses vetores B_i dependam da solução do sistema obtida usando vetores B_j anteriores. Nesse caso, é mais eficaz obter, primeiro, a fatoração LU de A e, depois, utilizar essa fatoração para resolver o sistema para cada novo vetor B.

Exemplo 3.23 Considere o sistema de equações lineares seguinte.

$$\begin{aligned} x + 2y + z &= k_1 \\ 2x + 3y + 3z &= k_2 \\ -3x + 10y + 2z &= k_3 \end{aligned} \quad \text{ou} \quad AX = B, \quad \text{com} \quad A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & 10 & 2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}$$

Digamos que queiramos resolver esse sistema três vezes, com B igual a, digamos, B_1, B_2 e B_3. Além disso, suponha que $B_1 = [1, 1, 1]^T$ e que

$$B_{j+1} = B_j + X_j \qquad (\text{para } j = 1, 2)$$

onde X_j é a solução de $AX = B_j$. Nesse caso, é mais eficaz obter, primeiro, a fatoração LU de A e, depois, utilizar essa fatoração para resolver o sistema para cada um dos vetores B. (Isso é feito nos Problemas 3.41 e 3.42.)

Problemas Resolvidos

Equações lineares, soluções, sistemas 2 × 2

3.1 Decida se as equações dadas são lineares.

(a) $5x + 7y - 8yz = 16$, (b) $x + \pi y + ez = \log 5$, (c) $3x + ky - 8z = 16$

(a) Não, porque o produto yz de duas incógnitas é de grau 2.

(b) Sim, porque π e $\log 5$ são constantes.

(c) Do jeito que está escrito, há quatro incógnitas, x, y, z, k e, por causa do termo ky, a equação não é linear. No entanto, considerando k constante, a equação é linear nas incógnitas x, y, z.

3.2 Decida se os vetores dados são soluções de $x_1 + 2x_2 - 4x_3 + 3x_4 = 15$:

(a) $u = (3, 2, 1, 4)$ e (b) $v = (1, 2, 4, 5)$.

(a) Substituindo, obtemos $3 + 2(2) - 4(1) + 3(4) = 15$, ou $15 = 15$; sim, é uma solução.

(b) Substituindo, obtemos $1 + 2(2) - 4(4) + 3(5) = 15$, ou $4 = 15$; não, não é uma solução.

3.3 Resolva (a) $ex = \pi$, (b) $3x - 4 - x = 2x + 3$, (c) $7 + 2x - 4 = 3x + 3 - x$

(a) Como $e \neq 0$, multiplicamos por $1/e$ para obter $x = \pi/e$.

(b) Reescrevendo a equação no formato padrão, obtemos $0x = 7$. A equação não tem solução.

(c) Reescrevendo a equação no formato padrão, obtemos $0x = 0$. Qualquer escalar k é uma solução.

3.4 Demonstre o Teorema 3.5. Considere a equação linear $ax = b$.

(i) Se $a \neq 0$, então $x = b/a$ é a única solução de $ax = b$.

(ii) Se $a = 0$, mas $b \neq 0$, então $ax = b$ não possui solução.

(iii) Se $a = 0$ e $b = 0$, então cada escalar k é uma solução de $ax = b$.

Suponha que $a \neq 0$. Então existe o escalar b/a. Substituir b/a em $ax = b$ fornece $a(b/a) = b$, ou $b = b$, portanto, b/a é uma solução. Por outro lado, digamos que x_0 seja uma solução de $ax = b$, ou seja, que $ax_0 = b$. Multiplicar ambos lados por $1/a$ fornece $x_0 = b/a$. Isso mostra que b/a é a única solução de $ax = b$ e demonstra (i). Suponha, agora, que $a = 0$. Então, dado qualquer escalar k, temos $ak = 0k = 0$. Se $b \neq 0$, então $ak \neq b$. e, por isso, k não é uma solução de $ax = b$, demonstrando (ii). Se $b = 0$, então $ak = b$, ou seja, qualquer escalar k é uma solução de $ax = b$, demonstrando (iii).

3.5 Resolva cada um dos sistemas seguintes.

(a) $\begin{aligned} 2x - 5y &= 11 \\ 3x + 4y &= 5 \end{aligned}$ (b) $\begin{aligned} 2x - 3y &= 8 \\ -6x + 9y &= 6 \end{aligned}$ (c) $\begin{aligned} 2x - 3y &= 8 \\ -4x + 6y &= -16 \end{aligned}$

(a) Eliminamos x das equações formando a nova equação $L = -3L_1 + 2L_2$. Isso fornece a equação

$$23y = -23 \qquad \text{e, então,} \qquad y = -1$$

Substituímos $y = -1$ em uma das equações originais, digamos, L_1, para obter

$$2x - 5(-1) = 11 \quad \text{ou} \quad 2x + 5 = 11 \quad \text{ou} \quad 2x = 6 \quad \text{ou} \quad x = 3$$

Assim, $x = 3$, $y = -1$ ou o par $u = (3, -1)$ é a única solução do sistema.

(b) Eliminamos x das equações formando a nova equação $L = 3L_1 + L_2$. Isso fornece a equação

$$0x + 0y = 30$$

Isso é uma equação degenerada com termo constante não nulo, portanto, essa equação e o sistema não têm solução. (Geometricamente, as retas correspondentes às equações são paralelas distintas.)

(c) Eliminamos x das equações formando a nova equação $L = 2L_1 + L_2$. Isso fornece a equação

$$0x + 0y = 0$$

Isso é uma equação degenerada com termo constante nulo, portanto, o sistema tem uma infinidade de soluções, que correspondem às soluções de qualquer uma das duas equações. (Geometricamente, as retas correspondentes às equações coincidem.)
Para encontrar a solução geral, tomamos $y = a$ e substituímos em L_1 para obter

$$2x - 3a = 8 \quad \text{ou} \quad 2x = 3a + 8 \quad \text{ou} \quad x = \tfrac{3}{2}a + 4$$

Assim, a solução geral é

$$x = \tfrac{3}{2}a + 4, \quad y = a \qquad \text{ou} \qquad u = \left(\tfrac{3}{2}a + 4,\ a\right)$$

com a um escalar qualquer.

3.6 Considere o sistema

$$\begin{aligned} x + ay &= 4 \\ ax + 9y &= b \end{aligned}$$

(a) Para quais valores de a o sistema tem solução única?

(b) Encontre os pares (a, b) de valores para os quais o sistema tem mais do que uma solução.

(a) Eliminamos x das equações formando a nova equação $L = -aL_1 + L_2$. Isso fornece a equação

$$(9 - a^2)y = b - 4a \tag{1}$$

O sistema tem solução única se, e só se, o coeficiente de y em (1) for não nulo, ou seja, se $9 - a^2 \neq 0$ ou se $a \neq \pm 3$.

(b) O sistema tem mais do que uma solução se ambos lados de (1) forem nulos. O lado esquerdo é nulo se $a = \pm 3$. Quando $a = 3$, o lado direito é nulo se $b - 12 = 0$, ou $b = 12$. Quando $a = -3$, o lado direito é nulo se $b + 12 = 0$, ou $b = -12$. Assim, $(3, 12)$ e $(-3, -12)$ são os pares para os quais o sistema tem mais do que uma solução.

Sistemas em forma triangular e escalonada

3.7 Encontre as variáveis pivôs e livres de cada um dos sistemas seguintes.

$$\begin{aligned} 2x_1 - 3x_2 - 6x_3 - 5x_4 + 2x_5 &= 7 \\ x_3 + 3x_4 - 7x_5 &= 6 \\ x_4 - 2x_5 &= 1 \end{aligned}$$
(a)

$$\begin{aligned} 2x - 6y + 7z &= 1 \\ 4y + 3z &= 8 \\ 2z &= 4 \end{aligned}$$
(b)

$$\begin{aligned} x + 2y - 3z &= 2 \\ 2x + 3y + z &= 4 \\ 3x + 4y + 5z &= 8 \end{aligned}$$
(c)

(a) Em forma escalonada, as incógnitas líderes são as variáveis pivôs e as outras são as variáveis livres. Aqui, x_1, x_3 e x_4, são as variáveis pivôs e x_2 e x_5 são as variáveis livres.

(b) As incógnitas líderes são x, y e z, portanto, são as variáveis pivôs. Não há variáveis livres (como em todo sistema triangular).

(c) O conceito de variável pivô e livre só se aplica a sistemas em forma escalonada.

3.8 Resolva o sistema triangular do Problema 3.7(b).

Por ser um sistema triangular, resolvemos por substituição para trás.

(i) A última equação dá $z = 2$.

(ii) Substituímos $z = 2$ na segunda equação e obtemos $4x + 6 = 8$, ou $y = \frac{1}{2}$.

(iii) Substituímos $z = 2$ e $y = \frac{1}{2}$ na primeira equação e obtemos

$$2x - 6\left(\frac{1}{2}\right) + 7(2) = 1 \qquad \text{ou} \qquad 2x + 11 = 1 \qquad \text{ou} \qquad x = -5$$

Assim, $x = -5$, $y = \frac{1}{2}$, $z = 2$ ou $u = (-5, \frac{1}{2}, 2)$ é a única solução do sistema.

3.9 Resolva o sistema escalonado do Problema 3.7(a).

Atribuímos parâmetros às variáveis livres, digamos, $x_2 = a$ e $x_5 = b$ e resolvemos nas variáveis pivôs por substituição para trás.

(i) Substituímos $x_5 = b$ na última equação e obtemos $x_4 - 2b = 1$, ou $x_4 = 2b + 1$.

(ii) Substituímos $x_5 = b$ e $x_4 = 2b + 1$ na segunda equação e obtemos

$$x_3 + 3(2b + 1) - 7b = 6 \qquad \text{ou} \qquad x_3 - b + 3 = 6 \qquad \text{ou} \qquad x_3 = b + 3$$

(iii) Substituímos $x_5 = b$, $x_4 = 2b + 1$, $x_3 = b + 3$ e $x_2 = a$ na primeira equação e obtemos

$$2x_1 - 3a - 6(b + 3) - 5(2b + 1) + 2b = 7 \qquad \text{ou} \qquad 2x_1 - 3a - 14b - 23 = 7$$

$$\text{ou} \qquad x_1 = \tfrac{3}{2}a + 7b + 15$$

Assim,

$$x_1 = \frac{3}{2}a + 7b + 15, \qquad x_2 = a, \qquad x_3 = b + 3, \qquad x_4 = 2b + 1, \quad x_5 = b$$

$$\text{ou} \qquad u = \left(\frac{3}{2}a + 7b + 15, \quad a, \quad b + 3, \quad 2b + 1, \quad b\right)$$

é a forma paramétrica da solução geral.

Alternativamente, resolvendo nas variáveis pivôs x_1, x_3, x_4 em termos das variáveis livres x_2 e x_5, obtemos a forma de variáveis livres da solução geral seguinte.

$$x_1 = \frac{3}{2}x_2 + 7x_5 + 15, \qquad x_3 = x_5 + 3, \qquad x_4 = 2x_5 + 1$$

3.10 Demonstre o Teorema 3.6. Considere o sistema (3.5) de equações lineares em forma escalonada de r equações com n incógnitas.

(i) Se $r = n$, então o sistema tem uma única solução.

(ii) Se $r < n$, então podemos atribuir arbitrariamente valores às $n - r$ variáveis livres e resolver de modo único nas r variáveis pivô, obtendo uma solução do sistema.

(i) Suponha que $r = n$. Então temos um sistema quadrado $AX = B$ em que a matriz de coeficientes é triangular (superior) com elementos diagonais não nulos. Assim, A é invertível. Pelo Teorema 3.10, o sistema tem uma solução única.

(ii) Atribuindo valores às $n - r$ variáveis livres, obtemos um sistema triangular nas variáveis pivôs, que, por (i), tem uma solução única.

Eliminação gaussiana

3.11 Resolva cada um dos sistemas seguintes.

$$\begin{array}{ccc} \begin{aligned} x+2y-4z&=-4\\ 2x+5y-9z&=-10\\ 3x-2y+3z&=11 \end{aligned} & \begin{aligned} x+2y-3z&=-1\\ -3x+y-2z&=-7\\ 5x+3y-4z&=2 \end{aligned} & \begin{aligned} x+2y-3z&=1\\ 2x+5y-8z&=4\\ 3x+8y-13z&=7 \end{aligned} \\ \text{(a)} & \text{(b)} & \text{(c)} \end{array}$$

Reduza cada sistema à forma triangular usando eliminação gaussiana.

(a) Aplicamos "Substituir L_2 por $-2L_1 + L_2$" e "Substituir L_3 por $-3L_1 + L_3$" para eliminar x da segunda e terceira equações e, depois, aplicamos "Substituir L_3 por $8L_2 + L_3$" para eliminar y da terceira equação. Essas operações fornecem

$$\begin{aligned} x+2y-4z&=-4\\ y-z&=-2\\ -8y+15z&=23 \end{aligned} \qquad \text{e, então} \qquad \begin{aligned} x+2y-4z&=-4\\ y-z&=-2\\ 7z&=7 \end{aligned}$$

O sistema está em forma triangular. Com substituição para trás, obtemos a solução única $u = (2, -1, 1)$.

(b) Eliminando x da segunda e terceira equações pelas operações "Substituir L_2 por $3L_1 + L_2$" e "Substituir L_3 por $-5L_1 + L_3$", obtemos o sistema equivalente

$$\begin{aligned} x+2y-3z&=-1\\ 7y-11z&=-10\\ -7y+11z&=7 \end{aligned}$$

A operação "Substituir L_3 por $L_2 + L_3$" fornece a equação degenerada de termo constante não nulo

$$0x + 0y + 0z = -3$$

Essa equação e, portanto, o sistema, não têm solução.

(c) Eliminando x da segunda e terceira equações pelas operações "Substituir L_2 por $-2L_1 + L_2$" e "Substituir L_3 por $-3L_1 + L_3$", obtemos o sistema equivalente

$$\begin{aligned} x+2y-3z&=1\\ y-2z&=2\\ 2y-4z&=4 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x+2y-3z&=1\\ y-2z&=2 \end{aligned}$$

(Omitimos a terceira equação por ser um múltiplo da segunda equação.) O sistema está em forma escalonada com variáveis pivôs x e y e variável livre z.

Para encontrar a forma paramétrica da solução geral, tomamos $z = a$ e resolvemos em x e y por substituição para trás. Substituindo $z = a$ na segunda equação, obtemos $y = 2 + 2a$. Substituindo $z = a$ e $y = 2 + 2a$ na primeira equação, obtemos

$$x + 2(2 + 2a) - 3a = 1 \qquad \text{ou} \qquad x + 4 + a = 1 \qquad \text{ou} \qquad x = -3 - a$$

Assim, a solução geral é

$$x = -3 - a,\ y = 2 + 2a,\ z = a \qquad \text{ou} \qquad u = (-3 - a,\ 2 + 2a,\ a),$$

onde a é um parâmetro.

3.12 Resolva cada um dos sistemas seguintes.

$$\begin{array}{cc} \begin{aligned} x_1-3x_2+2x_3-x_4+2x_5&=2\\ 3x_1-9x_2+7x_3-x_4+3x_5&=7\\ 2x_1-6x_2+7x_3+4x_4-5x_5&=7 \end{aligned} & \begin{aligned} x_1+2x_2-3x_3+4x_4&=2\\ 2x_1+5x_2-2x_3+x_4&=1\\ 5x_1+12x_2-7x_3+6x_4&=3 \end{aligned} \\ \text{(a)} & \text{(b)} \end{array}$$

Reduza cada sistema à forma escalonada usando eliminação gaussiana.

(a) Aplicamos "Substituir L_2 por $-3L_1 + L_2$" e "Substituir L_3 por $-2L_1 + L_3$" para eliminar x da segunda e terceira equações. Isso fornece

$$\begin{aligned} x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\ x_3 + 2x_4 - 3x_5 &= 1 \\ 3x_3 + 6x_4 - 9x_5 &= 3 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\ x_3 + 2x_4 - 3x_5 &= 1 \end{aligned}$$

(Omitimos L_3 por ser um múltiplo de L_2.) O sistema está em forma escalonada com variáveis pivôs x_1 e x_3 e variáveis livres x_2, x_4, x_5.

Para encontrar a forma paramétrica da solução geral, tomamos $x_2 = a$, $x_4 = b$, $x_5 = c$, em que a, b e c são parâmetros. Com substituição para trás, obtemos $x_3 = 1 - 2b + 3c$ e $x_1 = 3a + 5b - 8c$. A solução geral é

$$x_1 = 3a + 5b - 8c,\ x_2 = a,\ x_3 = 1 - 2b + 3c,\ x_4 = b,\ x_5 = c$$

ou, equivalentemente, $u = (3a + 5b - 8c, a, 1 - 2b + 3c, b, c)$.

(b) Eliminando x_1 da segunda e terceira equações pelas operações "Substituir L_2 por $-2L_1 + L_2$" e "Substituir L_3 por $-5L_1 + L_3$", obtemos o sistema equivalente

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 4x_4 &= 2 \\ x_2 + 4x_3 - 7x_4 &= -3 \\ 2x_2 + 8x_3 - 14x_4 &= -7 \end{aligned}$$

A operação "Substituir L_3 por $-2L_2 + L_3$" fornece a equação degenerada $0 = -1$. Assim, o sistema, não tem solução (embora tenha mais incógnitas do que equações).

3.13 Resolva usando o formato condensado.

$$\begin{aligned} 2y + 3z &= 3 \\ x + y + z &= 4 \\ 4x + 8y - 3z &= 35 \end{aligned}$$

O formato condensado é o seguinte.

	Número	Equação	Operação
(2)	($\not{1}$)	$2y + 3z = 3$	$L_1 \leftrightarrow L_2$
(1)	($\not{2}$)	$x + y + z = 4$	$L_1 \leftrightarrow L_2$
	(3)	$4x + 8y - 3z = 35$	
	(3′)	$4y - 7z = 19$	Substituir L_3 por $-4L_1 + L_3$
	(3″)	$-13z = 13$	Substituir L_3 por $-2L_2 + L_3$

Aqui, (1), (2) e (3″) formam um sistema triangular. (Enfatizamos que a troca de L_1 com L_2 foi efetuada simplesmente trocando os números das equações.)

Usar substituição para trás no sistema triangular fornece $z = -1$ em L_3, $y = 3$ em L_2 e $x = 2$ em L_1. Assim, a única solução do sistema é $x = 2$, $y = 3$, $z = -1$, ou o terno $u = (2, 3, -1)$.

3.14 Considere o sistema

$$\begin{aligned} x + 2y + z &= 3 \\ ay + 5z &= 10 \\ 2x + 7y + az &= b \end{aligned}$$

(a) Encontre os valores de a que fazem com que o sistema tenha solução única.

(b) Encontre os pares de valores (a, b) que fazem com que o sistema tenha mais do que uma solução.

Reduzimos o sistema à forma escalonada, ou seja, eliminamos x da terceira equação pela operação "Substituir L_3 por $-2L_1 + L_3$" e depois eliminamos y da terceira equação pela operação

"Substituir L_3 por $-3L_2 + aL_3$", obtendo

$$\begin{aligned} x + 2y \quad\ + \ z &= 3 \\ ay \quad + 5z &= 10 \\ 3y + (a-2)z &= b - 6 \end{aligned} \qquad \text{e, então,} \qquad \begin{aligned} x + 2y + \ z &= 3 \\ ay + 5z &= 10 \\ (a^2 - 2a - 15)z &= ab - 6a - 30 \end{aligned}$$

Examinemos a última equação $(a^2 - 2a - 15)z = ab - 6a - 30$.

(a) O sistema tem solução única se, e só se, o coeficiente de z não é zero, ou seja, se

$$a^2 - 2a - 15 = (a-5)(a+3) \neq 0 \qquad \text{ou} \qquad a \neq 5 \quad \text{e} \quad a \neq -3.$$

(b) O sistema tem mais do que uma solução se ambos lados são nulos. O lado esquerdo é nulo se $a = 5$ ou $a = -3$. Se $a = 5$, o lado direito é zero se $5b - 60 = 0$, ou $b = 12$. Se $a = -3$, o lado direito é zero se $-3b - 12 = 0$, ou $b = -4$. Assim, (5, 12) e (–3, –4) são os pares com os quais o sistema tem mais do que uma solução.

Matrizes escalonadas, equivalência por linhas, forma canônica por linhas

3.15 Reduza à forma escalonada as matrizes

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{bmatrix}, \qquad \text{(b)} \quad B = \begin{bmatrix} -4 & 1 & -6 \\ 1 & 2 & -5 \\ 6 & 3 & -4 \end{bmatrix}$$

(a) Usamos $a_{11} = 1$ como um pivô para obter zeros abaixo de a_{11}, ou seja, usamos as operações com as linhas "Substituir R_2 por $-2R_1 + R_2$" e "Substituir R_3 por $-3R_1 + R_3$". Depois usamos $a_{23} = 4$ como um pivô para obter zeros abaixo de a_{23}, ou seja, usamos a operação com as linhas "Substituir R_3 por $-3R_1 + R_3$". Essas operações fornecem

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Agora a matriz está em forma escalonada.

(b) Os cálculos manuais costumam ser mais simples se o elemento pivô for igual a 1. Por isso, trocamos R_1 com R_2. Em seguida, aplicamos as operações "Substituir R_2 por $4R_1 + R_2$" e "Substituir R_3 por $-6R_2 + R_3$" e, depois, "Substituir R_3 por $R_2 + R_3$". Essas operações fornecem

$$B \sim \begin{bmatrix} 1 & 2 & -5 \\ -4 & 1 & -6 \\ 6 & 3 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & -9 & 26 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{bmatrix}$$

Agora a matriz está em forma escalonada.

3.16 Descreva o algoritmo de redução por *pivotamento*. Também descreva as vantagens, se houver, da utilização desse algoritmo.

O algoritmo de redução se torna um algoritmo de pivotamento se escolhermos como pivô a_{1j_1} a entrada na coluna j de maior valor absoluto e se usarmos a operação com a linha

$$(-a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$$

A principal vantagem do algoritmo por pivotamento é que essa operação com as linhas envolve a divisão pelo pivô atual a_{1j_1} e, no computador, os erros de arredondamento podem ser substancialmente reduzidos se dividirmos por um número com o maior valor absoluto possível.

3.17 Seja $A = \begin{bmatrix} 2 & -2 & 2 & 1 \\ -3 & 6 & 0 & -1 \\ 1 & -7 & 10 & 2 \end{bmatrix}$. Reduza A à forma escalonada usando o algoritmo por pivotamento.

Inicialmente trocamos R_1 com R_2 para que possamos usar –3 como o pivô e, então, aplicamos as operações "Substituir R_2 por $\frac{2}{3}R_1 + R_2$" e "Substituir R_3 por $\frac{1}{3}R_1 + R_3$". Essas operações fornecem

$$A \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 2 & -2 & 2 & 1 \\ 1 & -7 & 10 & 2 \end{bmatrix} \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & 2 & 2 & \frac{1}{3} \\ 0 & -5 & 10 & \frac{5}{3} \end{bmatrix}$$

Agora trocamos R_2 com R_3 para que possamos usar –5 como o pivô e, então, aplicamos as operações "Substituir R_3 por $\frac{2}{5}R_2 + R_3$". Obtemos

$$A \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 2 & 2 & \frac{1}{3} \end{bmatrix} \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 0 & 6 & 1 \end{bmatrix}$$

Reduzimos a matriz à forma escalonada utilizando pivotamento.

3.18 Reduza cada uma das matrizes seguintes à forma canônica por linhas.

$$\text{(a)} \quad A = \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 8 & 8 & -1 & 26 & 23 \end{bmatrix}, \qquad \text{(b)} \quad B = \begin{bmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{bmatrix}$$

(a) Inicialmente reduzimos A à forma escalonada aplicando as operações "Substituir R_2 por $-2R_1 + R_2$" e "Substituir R_3 por $-4R_1 + R_3$" e, depois, "Substituir R_3 por $-R_2 + R_3$". Essas operações fornecem

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 3 & 2 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 4 & 2 \end{bmatrix}$$

Agora substituímos para trás na matriz escalonada para obter a forma canônica por linhas de A. Mais precisamente, primeiro multiplicamos R_3 por $\frac{1}{4}$ para obter o pivô $a_{34} = 1$ e, então, aplicamos as operações "Substituir R_2 por $2R_3 + R_2$" e "Substituir R_1 por $-6R_3 + R_1$". Essas operações fornecem

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

Agora multiplicamos R_2 por $\frac{1}{3}$, obtendo o pivô $a_{23} = 1$ e, então, aplicamos "Substituir R_1 por $R_2 + R_1$", obtendo

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

Finalmente, multiplicamos R_1 por $\frac{1}{2}$, obtendo o pivô $a_{11} = 1$. Assim, obtemos a forma canônica por linhas seguinte de A.

$$A \sim \begin{bmatrix} 1 & 1 & 0 & 0 & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

(b) Como B está em forma escalonada, substituímos para trás para obter

$$B \sim \begin{bmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & -9 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & -9 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A última matriz, que é a matriz identidade I, é a forma canônica por linhas de B. (Isso era de se esperar, porque B é invertível e sua forma canônica por linhas deve ser I.)

3.19 Descreva o algoritmo de eliminação de Gauss-Jordan, que também reduz uma matriz A arbitrária a sua forma canônica por linhas.

De certa forma, o algoritmo de Gauss-Jordan é análogo ao algoritmo de eliminação gaussiana, exceto que, nele, cada pivô é utilizado para colocar zeros tanto acima quanto abaixo do pivô, e não só abaixo, antes de passar para o próximo pivô. Além disso, numa variação desse algoritmo começamos *normalizando* cada linha, ou seja, obtemos um pivô unitário, antes de utilizá-lo para produzir os zeros nas outras linhas, em vez de *normalizar* no final do algoritmo.

3.20 Seja $A = \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 1 & 1 & 4 & -1 & 3 \\ 2 & 5 & 9 & -2 & 8 \end{bmatrix}$. Use Gauss-Jordan para encontrar a forma canônica por linhas de A.

Usamos $a_{11} = 1$ como um pivô para obter zeros abaixo de a_{11} aplicando as operações "Substituir R_2 por $-R_1 + R_2$" e "Substituir R_3 por $-2R_1 + R_3$". Isso fornece

$$A \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 9 & 3 & -4 & 4 \end{bmatrix}$$

Multiplicamos R_2 por $\frac{1}{3}$, obtendo o pivô $a_{22} = 1$ e, então, obtemos zeros abaixo e acima de a_{22} aplicando as operações "Substituir R_3 por $-9R_2 + R_3$" e "Substituir R_1 por $2R_2 + R_1$". Essas operações fornecem

$$A \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 9 & 3 & -4 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 2 & 1 \end{bmatrix}$$

Finalmente, multiplicamos R_3 por $\frac{1}{2}$ para obter o pivô $a_{34} = 1$ e, então, obtemos zeros abaixo e acima de a_{34} aplicando as operações "Substituir R_2 por $\frac{2}{3}R_3 + R_2$" e "Substituir R_1 por $\frac{1}{3}R_3 + R_1$". Essas operações fornecem

$$A \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & 0 & \frac{17}{6} \\ 0 & 1 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

que é a forma canônica por linhas de A.

Sistemas de equações lineares em forma matricial

3.21 Encontre a matriz aumentada M e a matriz de coeficientes A do sistema seguinte.

$$\begin{aligned} x + 2y - 3z &= 4 \\ 3y - 4z + 7x &= 5 \\ 6z + 8x - 9y &= 1 \end{aligned}$$

Inicialmente, alinhamos as incógnitas do sistema e, depois, usamos o sistema alinhado para obter M e A. Temos

$$\begin{aligned} x + 2y - 3z &= 4 \\ 7x + 3y - 4z &= 5\,; \\ 8x - 9y + 6z &= 1 \end{aligned} \qquad \text{então} \qquad M = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 7 & 3 & -4 & 5 \\ 8 & -9 & 6 & 1 \end{bmatrix} \qquad \text{e} \qquad A = \begin{bmatrix} 1 & 2 & -3 \\ 7 & 3 & -4 \\ 8 & -9 & 6 \end{bmatrix}$$

3.22 Resolva cada um dos sistemas seguintes utilizando sua matriz aumentada M.

$$\begin{aligned} x + 2y - \;\; z &= \;\; 3 \\ x + 3y + \;\; z &= \;\; 5 \\ 3x + 8y + 4z &= 17 \end{aligned} \qquad \begin{aligned} x - 2y + 4z &= 2 \\ 2x - 3y + 5z &= 3 \\ 3x - 4y + 6z &= 7 \end{aligned} \qquad \begin{aligned} x + \;\; y + 3z &= 1 \\ 2x + 3y - \;\; z &= 3 \\ 5x + 7y + \;\; z &= 7 \end{aligned}$$

(a) (b) (c)

(a) Reduzimos a matriz aumentada M à forma escalonada como segue.

$$M = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 1 & 3 & 1 & 5 \\ 3 & 8 & 4 & 17 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 7 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 3 & 4 \end{bmatrix}$$

Agora escrevemos o sistema triangular correspondente

$$\begin{aligned} x + 2y - z &= 3 \\ y + 2z &= 2 \\ 3z &= 4 \end{aligned}$$

e substituímos para trás para obter a solução única

$$x = \tfrac{17}{3},\ \ y = -\tfrac{2}{3},\ \ z = \tfrac{4}{3} \qquad \text{ou} \qquad u = \left(\tfrac{17}{3}, -\tfrac{2}{3}, \tfrac{4}{3}\right)$$

Alternativamente, reduzimos a forma escalonada de M à forma canônica por linhas, obtendo

$$M \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{13}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix}$$

Isso também corresponde à solução encontrada antes.

(b) Inicialmente reduzimos a matriz aumentada M à forma escalonada como segue.

$$M = \begin{bmatrix} 1 & -2 & 4 & 2 \\ 2 & -3 & 5 & 3 \\ 3 & -4 & 6 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 2 & -6 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

A terceira linha corresponde à equação degenerada $0x + 0y + 0z = 3$, que não tem solução. Assim, “NÃO CONTINUE”. O sistema original também não tem solução. (Observe que a forma escalonada indica se o sistema tem ou não tem solução.)

(c) Reduzimos a matriz aumentada M à forma escalonada e, depois, à forma canônica por linhas, como segue.

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 \\ 2 & 3 & -1 & 3 \\ 5 & 7 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 1 \\ 0 & 1 & -7 & 1 \\ 0 & 2 & -14 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 10 & 0 \\ 0 & 1 & -7 & 1 \end{bmatrix}$$

(Omitimos a terceira linha da segunda matriz por ser um múltiplo da segunda linha e acarretar uma linha de zeros.) Escrevemos o sistema correspondente à forma canônica por linhas de M e depois transferimos as variáveis livres para o outro lado para obter a forma de variáveis livres da solução

$$\begin{aligned} x + 10z &= 0 \\ y - \ \ 7z &= 1 \end{aligned} \qquad \text{e} \qquad \begin{aligned} x &= -10z \\ y &= 1 + 7z \end{aligned}$$

Aqui, z é a única variável livre. A solução paramétrica, usando $z = a$, é dada a seguir.

$$x = -10a,\ \ y = 1 + 7a,\ \ z = a \qquad \text{ou} \qquad u = (-10a,\ 1 + 7a,\ a)$$

3.23 Resolva o sistema seguinte utilizando sua matriz aumentada M.

$$\begin{aligned} x_1 + 2x_2 - 3x_3 - 2x_4 + 4x_5 &= 1 \\ 2x_1 + 5x_2 - 8x_3 - \ \ x_4 + 6x_5 &= 4 \\ x_1 + 4x_2 - 7x_3 + 5x_4 + 2x_5 &= 8 \end{aligned}$$

Reduzimos a matriz aumentada M à forma escalonada e, depois, à forma canônica por linhas.

$$M = \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 2 & 5 & -8 & -1 & 6 & 4 \\ 1 & 4 & -7 & 5 & 2 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 2 & -4 & 7 & -2 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -3 & 0 & 8 & 7 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 0 & 24 & 21 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

Escrevemos o sistema correspondente à forma canônica por linhas de M e depois transferimos as variáveis livres para o outro lado para obter a forma de variáveis livres da solução

$$\begin{aligned} x_1 \ + x_3 + \quad\ 24x_5 &= \ 21 \\ x_2 - 2x_3 - \quad\ \ 8x_5 &= -7 \\ x_4 + 2x_5 &= \ \ 3 \end{aligned} \qquad \text{e} \qquad \begin{aligned} x_1 &= 21 - x_3 - 24x_5 \\ x_2 &= -7 + 2x_3 + 8x_5 \\ x_4 &= 3 - 2x_5 \end{aligned}$$

Aqui, x_1, x_2 e x_4 são as variáveis pivôs e x_3 e x_5 são as variáveis livres. Lembre que a forma paramétrica da solução pode ser obtida a partir da forma de variáveis livres da solução simplesmente igualando as variáveis livres a parâmetros, digamos, $x_3 = a$ e $x_5 = b$. Assim, obtemos

$$x_1 = 21 - a - 24b,\ x_2 = -7 + 2a + 8b,\ x_3 = a,\ x_4 = 3 - 2b,\ x_5 = b$$

ou

$$u = (21 - a - 24b, -7 + 2a + 8b, a, 3 - 2b, b)$$

que é outra forma da solução.

Combinações lineares, sistemas homogêneos

3.24 Escreva v como combinação linear dos vetores u_1, u_2 e u_3 dados a seguir.

(a) $v = (3, 10, 7)$ e $u_1 = (1, 3, -2), u_2 = (1, 4, 2), u_3 = (2, 8, 1)$;

(b) $v = (2, 7, 10)$ e $u_1 = (1, 2, 3),\ u_2 = (1, 3, 5),\ u_3 = (1, 5, 9)$;

(c) $v = (1, 5, 4)$ e $u_1 = (1, 3, -2),\ u_2 = (2, 7, -1),\ u_3 = (1, 6, 7)$.

Formamos o sistema de equações lineares equivalente escrevendo $v = xu_1 + yu_2 + zu_3$ ou, então, usamos a matriz aumentada M do sistema equivalente, em que $M = [u_1, u_2, u_3, v]$. (Aqui, u_1, u_2, u_3 e v são as colunas de M.)

(a) A equação vetorial $v = xu_1 + yu_2 + zu_3$ com os vetores dados é dada por

$$\begin{bmatrix} 3 \\ 10 \\ 7 \end{bmatrix} = x \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + y \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + z \begin{bmatrix} 2 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ 3x + 4y + 8z \\ -2x + 2y + z \end{bmatrix}$$

Igualando as entradas correspondentes, formamos o sistema de equações lineares equivalente que reduzimos à forma escalonada, como segue.

$$\begin{aligned} x + \ \ y + 2z &= \ \ 3 \\ 3x + 4y + 8z &= 10 \\ -2x + 2y + \ \ z &= \ \ 7 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + \ \ y + 2z &= \ \ 3 \\ y + 2z &= \ \ 1 \\ 4y + 5z &= 13 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + y + 2z &= 3 \\ y + 2z &= 1 \\ -3z &= 9 \end{aligned}$$

O sistema está em forma triangular. Substituindo para trás, obtemos $x = 2, y = 7, z = -3$. Assim, $v = 2u_1 + 7u_2 - 3u_3$.
Alternativamente, formamos a matriz aumentada $M = [u_1, u_2, u_3, v]$ do sistema equivalente e reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 3 & 4 & 8 & 10 \\ -2 & 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 4 & 5 & 13 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -3 & 9 \end{bmatrix}$$

A última matriz corresponde a um sistema triangular que tem uma solução única. Substituindo para trás, obtemos a solução $x = 2, y = 7, z = -3$. Assim, $v = 2u_1 + 7u_2 - 3u_3$.

(b) Formamos a matriz aumentada $M = [u_1, u_2, u_3, v]$ do sistema equivalente e reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & 5 & 7 \\ 3 & 5 & 9 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 2 & 6 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

A terceira linha corresponde à equação degenerada $0x + 0y + 0z = -2$, que não tem solução. Assim, o sistema também não tem solução, e v não pode ser escrito como uma combinação linear de u_1, u_2, u_3.

(c) Formamos a matriz aumentada $M = [u_1, u_2, u_3, v]$ do sistema equivalente e reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 7 & 6 & 5 \\ -2 & -1 & 7 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 3 & 9 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A última matriz corresponde a um sistema com uma variável livre z.

$$\begin{aligned} x + 2y + \ z &= 1 \\ y + 3z &= 2 \end{aligned}$$

Assim, v pode ser escrito de várias maneiras como uma combinação linear de u_1, u_2, u_3. Por exemplo, tomando a variável livre como $z = 1$ e substituindo para trás, obtemos $y = -2$ e $x = 2$. Assim, $v = 2u_1 - 2u_2 + u_3$.

3.25 Sejam $u_1 = (1, 2, 4)$, $u_2 = (2, -3, 1)$, $u_3 = (2, 1, -1)$ vetores de $\mathbf{R}^3$. Mostre que u_1, u_2, u_3 são ortogonais e escreva v como uma combinação linear de u_1, u_2, u_3 se (a) $v = (7, 16, 6)$, (b) $v = (3, 5, 2)$.

Tomamos o produto escalar dos pares de vetores e obtemos

$$u_1 \cdot u_2 = 2 - 6 + 4 = 0, \quad u_1 \cdot u_3 = 2 + 2 - 4 = 0, \quad u_2 \cdot u_3 = 4 - 3 - 1 = 0$$

Assim, os três vetores de $\mathbf{R}^3$ são ortogonais e, portanto, podemos usar coeficientes de Fourier, ou seja, $v = xu_1 + yu_2 + zu_3$, com

$$x = \frac{v \cdot u_1}{u_1 \cdot u_1}, \qquad y = \frac{v \cdot u_2}{u_2 \cdot u_2}, \qquad z = \frac{v \cdot u_3}{u_3 \cdot u_3}$$

(a) Temos

$$x = \frac{7 + 32 + 24}{1 + 4 + 16} = \frac{63}{21} = 3, \qquad y = \frac{14 - 48 + 6}{4 + 9 + 1} = \frac{-28}{14} = -2, \qquad z = \frac{14 + 16 - 6}{4 + 1 + 1} = \frac{24}{6} = 4$$

Assim, $v = 3u_1 - 2u_2 + 4u_3$.

(b) Temos

$$x = \frac{3 + 10 + 8}{1 + 4 + 16} = \frac{21}{21} = 1, \qquad y = \frac{6 - 15 + 2}{4 + 9 + 1} = \frac{-7}{14} = -\frac{1}{2}, \qquad z = \frac{6 + 5 - 2}{4 + 1 + 1} = \frac{9}{6} = \frac{3}{2}$$

Assim, $v = u_1 - \frac{1}{2}u_2 + \frac{3}{2}u_3$.

3.26 Encontre a dimensão e uma base do conjunto solução W de cada um dos sistemas homogêneos seguintes.

$$\begin{aligned} 2x_1 + \ 4x_2 - \ 5x_3 + 3x_4 &= 0 \\ 3x_1 + \ 6x_2 - \ 7x_3 + 4x_4 &= 0 \\ 5x_1 + 10x_2 - 11x_3 + 6x_4 &= 0 \end{aligned} \qquad \begin{aligned} x - 2y - 3z &= 0 \\ 2x + \ y + 3z &= 0 \\ 3x - 4y - 2z &= 0 \end{aligned}$$

(a) (b)

(a) Reduzimos o sistema à forma escalonada usando as operações "Substituir L_2 por $-3L_1 + 2L_2$", "Substituir L_3 por $-5L_1 + 2L_3$" e, depois, "Substituir L_3 por $-2L_2 + L_3$". Essas operações fornecem

$$\begin{aligned} 2x_1 + 4x_2 - 5x_3 + 3x_4 &= 0 \\ x_3 - \ x_4 &= 0 \\ 3x_3 - 3x_4 &= 0 \end{aligned} \qquad \text{e} \qquad \begin{aligned} 2x_1 + 4x_2 - 5x_3 + 3x_4 &= 0 \\ x_3 - \ x_4 &= 0 \end{aligned}$$

O sistema escalonado tem as duas variáveis livres x_2 e x_4, portanto dim $W = 2$. Uma base de W pode ser obtida como segue.

(1) Tomando $x_2 = 1$, $x_4 = 0$, substituímos para trás para obter $x_3 = 0$ e, então, $x_1 = -2$. Assim, $u_1 = (-2, 1, 0, 0)$.

(2) Tomando $x_2 = 0$, $x_4 = 1$, substituímos para trás para obter $x_3 = 1$ e, então, $x_1 = 1$. Assim, $u_2 = (1, 0, 1, 1)$.

(b) Reduzindo o sistema à forma escalonada, obtemos

$$\begin{aligned} x - 2y - 3z &= 0 \\ 5y + 9z &= 0 \\ 2y + 7z &= 0 \end{aligned} \qquad \text{e} \qquad \begin{aligned} x - 2y - 3z &= 0 \\ 5y + 9z &= 0 \\ 17z &= 0 \end{aligned}$$

Não há variáveis livres (o sistema está em forma triangular). Logo, dim $W = 0$ e W não possui base. Mais precisamente, W consiste somente na solução nula, ou seja, $W = \{0\}$.

3.27 Utilizando notação matricial, encontre a dimensão e uma base do conjunto solução W do sistema homogêneo seguinte.

$$\begin{aligned} x_1 + 2x_2 + \ 3x_3 - 2x_4 + \ 4x_5 &= 0 \\ 2x_1 + 4x_2 + \ 8x_3 + \ x_4 + \ 9x_5 &= 0 \\ 3x_1 + 6x_2 + 13x_3 + 4x_4 + 14x_5 &= 0 \end{aligned}$$

Mostre como a base dá a forma paramétrica da solução geral do sistema.

Um sistema homogêneo é representado por sua matriz de coeficientes A em vez da matriz aumentada M, porque a última coluna da matriz aumentada M é nula e permanece nula ao longo de qualquer algoritmo de redução.

Reduzimos a matriz de coeficientes A à forma escalonada, obtendo

$$A = \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 2 & 4 & 8 & 1 & 9 \\ 3 & 6 & 13 & 4 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 2 & 5 & 1 \\ 0 & 0 & 4 & 10 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 2 & 5 & 1 \end{bmatrix}$$

(A terceira linha da segunda matriz foi suprimida, por ser um múltiplo da segunda linha e resultar numa linha nula.) Agora podemos continuar de duas maneiras.

(a) Escrevemos o sistema homogêneo correspondente em forma escalonada.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - 2x_4 + 4x_5 &= 0 \\ 2x_3 + 5x_4 + x_5 &= 0 \end{aligned}$$

O sistema escalonado tem as três variáveis livres x_2, x_4 e x_5, portanto dim $W = 3$. Uma base $[u_1, u_2, u_3]$ de W pode ser obtida como segue.

(1) Tomando $x_2 = 1$, $x_4 = 0$, $x_5 = 0$, substituímos para trás para obter $x_3 = 0$ e, então, $x_1 = -2$. Assim,

$$u_1 = (-2, 1, 0, 0, 0).$$

(2) Tomando $x_2 = 0$, $x_4 = 1$, $x_5 = 0$, substituímos para trás para obter $x_3 = -\frac{5}{2}$, e, então, $x_1 = \frac{19}{2}$. Assim,

$$u_2 = (\tfrac{19}{2}, 0, -\tfrac{5}{2}, 1, 0).$$

(3) Tomando $x_2 = 0$, $x_4 = 0$, $x_5 = 1$, substituímos para trás para obter $x_3 = -\frac{1}{2}$, e, então, $x_1 = -\frac{5}{2}$. Assim,

$$u_3 = (-\tfrac{5}{2}, 0, -\tfrac{1}{2}, 0, 1).$$

[Poderíamos evitar frações na base escolhendo $x_4 = 2$ em (2) e $x_5 = 2$ em (3), que produzem múltiplos de u_2 e u_3.] Obtemos a forma paramétrica da solução geral como uma combinação linear dos vetores da base utilizando parâmetros x, y e z.

$$xu_1 + yu_2 + zu_3 = (-2x + \tfrac{19}{2}y - \tfrac{5}{2}z,\ \ x,\ \ -\tfrac{5}{2}y - \tfrac{1}{2}z,\ \ y,\ \ z)$$

(b) Reduzimos a forma escalonada de A à forma canônica, como segue.

$$A \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -\frac{19}{2} & \frac{5}{2} \\ 0 & 0 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix}$$

Agora escrevemos a correspondente solução de variáveis livres

$$x_1 = -2x_2 + \frac{19}{2}x_4 - \frac{5}{2}x_5$$

$$x_3 = -\frac{5}{2}x_4 - \frac{1}{2}x_5$$

Usando essas equações para as variáveis pivôs x_1 e x_3, repetimos o processo da opção (a) para obter uma base $[u_1, u_2, u_3]$ de W. Ou seja, tomamos $x_2 = 1$, $x_4 = 0$, $x_5 = 0$ para obter u_1; tomamos $x_2 = 0$, $x_4 = 1$, $x_5 = 0$ para obter u_2 e tomamos $x_2 = 0$, $x_4 = 0$, $x_5 = 1$ para obter u_3.

3.28 Demonstre o Teorema 3.15. Sejam v_0 uma solução particular de $AX = B$ e W o conjunto solução de $AX = 0$. Então $U = v_0 + W = \{v_0 + w :\ w \in W\}$ é o conjunto solução de $AX = B$.

Seja w uma solução de $AX = 0$. Então

$$A(v_0 + w) = Av_0 + Aw = B + 0 = B$$

Assim, a soma $v_0 + w$ é uma solução de $AX = B$. Por outro lado, seja v uma solução de $AX = B$. Então

$$A(v - v_0) = Av - Av_0 = B - B = 0$$

Portanto, $v - v_0$ pertence a W. Como $v = v_0 + (v - v_0)$, vemos que toda solução de $AX = B$ pode ser dada pela soma de uma solução de $AX = 0$ com uma solução de $AX = B$. O teorema está demonstrado.

Matrizes elementares, aplicações

3.29 Sejam e_1, e_2, e_3, respectivamente, as operações elementares com as linhas

"Trocar as linhas R_1 e R_2 entre si", "Substituir R_3 por $7R_3$" e "Substituir R_2 por $-3R_1 + R_2$"

Encontre as matrizes elementares E_1, E_2, E_3 de ordem 3 correspondentes.

Aplicando cada operação à matriz identidade I_3 de ordem 3, obtemos

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.30 Considere as operações elementares com as linhas dadas no Problema 3.29.

(a) Descreva as operações inversas e_1^{-1}, e_2^{-1}, e_3^{-1}.

(b) Encontre as matrizes elementares E'_1, E'_2, E'_3 de ordem 3 correspondentes.

(c) Qual é a relação entre as matrizes E'_1, E'_2, E'_3 e as matrizes E_1, E_2, E_3?

(a) As inversas de e_1, e_2, e_3 são, respectivamente,

"Trocar as linhas R_1 e R_2 entre si", "Substituir R_3 por $\frac{1}{7}R_3$" e "Substituir R_2 por $3R_1 + R_2$"

(b) Aplicando cada operação inversa à matriz identidade I_3 de ordem 3, obtemos

$$E'_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E'_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}, \qquad E'_3 = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c) As matrizes E'_1, E'_2, E'_3 são, respectivamente, as inversas das matrizes E_1, E_2, E_3.

3.31 Escreva cada uma das matrizes seguintes como um produto de matrizes elementares.

(a) $A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 8 \\ -3 & -1 & 2 \end{bmatrix}$

Para escrever uma matriz M como um produto de matrizes elementares utilizamos os três passos seguintes.

Passo 1 Reduzimos M à matriz identidade I, fazendo um registro das operações utilizadas.

Passo 2 Escrevemos as operações inversas das registradas.

Passo 3 Escrevemos M como o produto das matrizes elementares correspondentes às operações inversas, que é o resultado procurado.

Se aparecer uma linha nula no Passo 1, então M não é equivalente por linhas à matriz identidade e M não pode ser escrita como um produto de matrizes elementares.

(a) (1) Temos

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

em que as operações com as linhas são, respectivamente,

"Substituir R_2 por $2R_1 + R_2$", "Substituir R_2 por $-\frac{1}{2}R_2$", "Substituir R_1 por $3R_2 + R_1$"

(2) As operações inversas são

"Substituir R_2 por $-2R_1 + R_2$", "Substituir R_2 por $-2R_2$", "Substituir R_1 por $-3R_2 + R_1$"

(3) $A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$

(b) (1) Temos

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

em que as operações com as linhas são, respectivamente,

"Substituir R_2 por $-4R_3 + R_2$", "Substituir R_1 por $-3R_3 + R_1$", "Substituir R_1 por $-2R_2 + R_1$"

(2) As operações inversas são

"Substituir R_2 por $4R_3 + R_2$", "Substituir R_1 por $3R_3 + R_1$", "Substituir R_1 por $2R_2 + R_1$"

(3) $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) (1) Começamos reduzindo C à forma escalonada. Temos

$$C = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 8 \\ -3 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 2 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

A forma escalonada de C tem uma linha nula. "PARE". A matriz C não pode ser reduzida à matriz identidade e C não pode ser escrita como um produto de matrizes elementares. (Em particular, observamos que C não possui inversa.)

3.32 Encontre a inversa de (a) $A = \begin{bmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{bmatrix}$.

(a) Formamos a matriz $M = [A, I]$ e a reduzimos à forma escalonada.

$$M = \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ -1 & -1 & 5 & 0 & 1 & 0 \\ 2 & 7 & -3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 3 & 5 & -2 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & -5 & -3 & 1 \end{array}\right]$$

Na forma escalonada, a metade da esquerda de M está em forma triangular, portanto, A tem uma inversa. Continuamos reduzindo M à forma canônica por linhas.

$$M \sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -9 & -6 & 2 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -16 & -11 & 3 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right]$$

A última matriz tem o formato $[I, A^{-1}]$, ou seja, A^{-1} é a metade da direita da última matriz. Assim,

$$A^{-1} = \begin{bmatrix} -16 & -11 & 3 \\ \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

(b) Formamos a matriz $M = [B, I]$ e a reduzimos à forma escalonada.

$$M = \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 1 & 5 & -1 & 0 & 1 & 0 \\ 3 & 13 & -6 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 0 \\ 0 & 4 & 6 & -3 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 \end{array}\right]$$

Na forma escalonada, M tem uma linha nula na metade da esquerda, portanto, B não pode ser reduzida à forma triangular. Por isso, B não tem uma inversa.

3.33 Mostre que cada matriz elementar E é invertível e que sua inversa é uma matriz elementar.

Seja E a matriz elementar correspondente à operação elementar e, ou seja, $e(I) = E$. Seja e' a operação inversa de e e E' a matriz elementar correspondente, ou seja, $e'(I) = E'$. Então

$$I = e'(e(I)) = e'(E) = E'E \qquad \text{e} \qquad I = e(e'(I)) = e(E') = EE'$$

Assim, E' é a inversa de E.

3.34 Demonstre o Teorema 3.16. Sejam e uma operação elementar com as linhas e E a matriz elementar quadrada de ordem m correspondente, ou seja, $e(I) = E$. Então $e(A) = EA$, para qualquer matriz A $m \times n$.

Seja R_i a i-ésima linha de A e escrevamos $A = [R_1, ..., R_m]$. Se B for um matriz para a qual está definido o produto AB, então $AB = [R_1B, ..., R_mB.]$. Também escrevemos

$$e_i = (0, \ldots, 0, \hat{1}, 0, \ldots, 0), \qquad \hat{} = i$$

Aqui, $\hat{} = i$ significa que a entrada 1 com circunflexo está na posição i. Mostra-se (Problema 2.46) que $e_iA = R_i$. Observe que $I = [e_1, e_2, ..., e_m]$ é a identidade de ordem m.

(i) Seja e a operação elementar "Trocar as linhas R_1 e R_2 entre si". Então, com $\hat{} = i$ e $\hat{\hat{}} = j$,

$$E = e(I) = [e_1, \ldots, \widehat{e_j}, \ldots, \widehat{\widehat{e_i}}, \ldots, e_m]$$

e

$$e(A) = [R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m]$$

Assim,

$$EA = [e_1A, \ldots, \widehat{e_jA}, \ldots, \widehat{\widehat{e_iA}}, \ldots, e_mA] = [R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m] = e(A)$$

(ii) Seja e a operação elementar "Substituir R_i por $kR_i\,(k \neq 0)$". Então, com $\hat{} = i$,

$$E = e(I) = [e_1, \ldots, \widehat{ke_i}, \ldots, e_m]$$

e

$$e(A) = [R_1, \ldots, \widehat{kR_i}, \ldots, R_m]$$

Assim,

$$EA = [e_1A, \ldots, \widehat{ke_iA}, \ldots, e_mA] = [R_1, \ldots, \widehat{kR_i}, \ldots, R_m] = e(A)$$

(iii) Seja e a operação elementar "Substituir R_i por $kR_j + R_i$". Então, com $\hat{} = i$,

$$E = e(I) = [e_1, \ \ldots, \ \widehat{ke_j + e_i}, \ \ldots, \ e_m]$$

e

$$e(A) = [R_1, \ \ldots, \ \widehat{kR_j + R_i}, \ \ldots, \ R_m]$$

Usando $(ke_j + e_i)A = k(e_jA) + e_iA = kR_j + R_i$, obtemos

$$\begin{aligned} EA &= [e_1A, \quad \ldots, \quad (ke_j + e_i)A, \quad \ldots, \quad e_mA] \\ &= [R_1, \quad \ldots, \quad \widehat{kR_j + R_i}, \quad \ldots, \quad R_m] = e(A) \end{aligned}$$

3.35 Demonstre o Teorema 3.17. Seja A uma matriz quadrada. Então as afirmativas seguintes são equivalentes.

(a) A é invertível (não singular).

(b) A é equivalente por linhas à matriz identidade I.

(c) A é um produto de matrizes elementares.

Suponha que A seja invertível e equivalente por linhas a uma matriz B em forma canônica. Então existem matrizes elementares $E_1, E_2, ..., E_s$ tais que $E_s \ldots E_2E_1A = B$. Como A e cada matriz elementar E_i são invertíveis, também B é invertível. No entanto, se $B \neq I$, então B teria uma linha nula e, portanto, não seria invertível. Logo $B = I$ e (a) implica (b).

Se valer (b), existem matrizes elementares E_1, E_2, ..., E_s tais que $E_s \ldots E_2E_1A = I$. Logo, $A = (E_s \ldots E_2E_1)^{-1} = E_1^{-1}E_2^{-1} \ldots, E_s^{-1}$. Como as E_i^{-1} também são matrizes elementares, (b) implica (c).
Se vale (c), então $A = E_1E_2 \ldots E_s$ e, como as E_i são matrizes invertíveis, seu produto A também é invertível. Logo, (c) implica (a). Dessa forma, demonstramos o teorema.

3.36 Demonstre o Teorema 3.18. Se $AB = I$, então $BA = I$ e, portanto, $B = A^{-1}$.

Suponha que A não seja invertível. Então A não é equivalente por linhas à matriz identidade I e, portanto, A é equivalente por linhas a uma matriz com uma linha nula. Em outras palavras, existem matrizes elementares E_1, ..., E_s tais que $E_s \ldots E_2E_1A$ tem uma linha nula. Segue que a matriz invertível $E_s \ldots E_2E_1AB = E_s \ldots E_2E_1$ tem uma linha nula. Como matrizes invertíveis não podem ter linhas nulas, isso mostra que A é invertível, com inversa A^{-1}. Então também

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}$$

3.37 Demonstre o Teorema 3.19. B é equivalente por linhas a A (e escrevemos $B \sim A$) se, e só se, existe uma matriz não singular P tal que $B = PA$.

Se $B \sim A$, então $B = e_s(\ldots(e_2(e_1(A)))\ldots) = E_s \ldots E_2E_1A = PA$, onde $P = E_s \ldots E_2E_1$ é não singular. Reciprocamente, suponha que $B = PA$, com P não singular. Pelo Teorema 3.17, P é um produto de matrizes elementares, portanto, B pode ser obtido a partir de A com uma sequência de operações elementares com as linhas, ou seja, $B \sim A$. Dessa forma, demonstramos o teorema.

3.38 Demonstre o Teorema 3.21. Cada matriz $m \times n$ A é equivalente a uma única matriz em blocos da forma $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, em que I_r é a matriz identidade de ordem r.

A demonstração é construtiva, no formato de algoritmo.

Passo 1 Reduza A à forma canônica por linhas, com entradas líderes não nulas $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$.

Passo 2 Troque entre si as colunas C_1 e C_{1j_1}, C_2 e $C_{2j_2}, \ldots, C_r$ e C_{jr}. Disso resulta uma matriz da forma $\left[\begin{array}{c|c} I_r & B \\ \hline 0 & 0 \end{array}\right]$, com entradas líderes não nulas $a_{11}, a_{22}, \ldots, a_{rr}$.

Passo 3 Utilize operações com as colunas e os pivôs a_{ii} para trocar cada entrada de B com um zero, ou seja, para $i = 1, 2, ..., r$ e $j = r + 1, r + 2, ..., n$ efetue a operação $-b_{ij}C_i + C_j \to C_j$.

A matriz final tem o formato $\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$ procurado.

Fatoração LU

3.39 Encontre a fatoração LU de (a) $A = \begin{bmatrix} 1 & -3 & 5 \\ 2 & -4 & 7 \\ -1 & -2 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 4 & -3 \\ 2 & 8 & 1 \\ -5 & -9 & 7 \end{bmatrix}$.

(a) Reduzimos A à forma triangular com as operações

"Substituir R_2 por $-2R_1 + R_2$", "Substituir R_3 por $R_1 + R_3$" e, então,

"Substituir R_3 por $\frac{5}{2}R_2 + R_3$"

Denotando a forma triangular por U, essas operações fornecem

$$A \sim \begin{bmatrix} 1 & -3 & 5 \\ 0 & 2 & -3 \\ 0 & -5 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & -\frac{3}{2} \end{bmatrix} = U \qquad \text{e} \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{5}{2} & 1 \end{bmatrix}$$

As entradas $2, -1, -\frac{5}{2}$ de L são os simétricos dos multiplicadores $-2, 1, \frac{5}{2}$ das operações com as linhas utilizadas. (Para conferir, multiplique L por U e obtenha $A = LU$.)

(b) Reduzimos B à forma triangular com as operações "Substituir R_2 por $-2R_1 + R_2$", "Substituir R_3 por $5R_1 + R_3$". Essas operações fornecem

$$B \sim \begin{bmatrix} 1 & 4 & -3 \\ 0 & 0 & 7 \\ 0 & 11 & -8 \end{bmatrix}.$$

Observe que a segunda diagonal é 0. Assim, B não pode ser levado à forma triangular sem utilizar troca de linhas. Por isso, B não tem fatoração LU. (Existe uma fatoração PLU desse tipo de matrizes, em que P é uma matriz de permutação, mas essa fatoração não está ao alcance de nossos métodos.)

3.40 Encontre a fatoração LDU da matriz A do Problema 3.39.

A fatoração LDU se refere ao caso em que L é uma matriz triangular inferior com entradas 1 na diagonal (como na fatoração LU), D é uma matriz diagonal e U é uma matriz triangular superior com entradas 1 na diagonal. Assim, simplesmente separamos as entradas da diagonal da matriz U da fatoração LU de A obtida no problema precedente e obtemos as matrizes D e U, como segue.

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{5}{2} & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -\frac{3}{2} \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & -3 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

3.41 Encontre a fatoração LU da matriz $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & -10 & 2 \end{bmatrix}$.

Reduzimos A à forma triangular com as operações

(1) "Substituir R_2 por $-2R_1 + R_2$", (2) "Substituir R_3 por $3R_1 + R_3$", (3) "Substituir R_3 por $-4R_2 + R_3$"

Denotando a forma triangular por U, essas operações fornecem

$$A \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & -4 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = U \qquad \text{e} \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{bmatrix}$$

As entradas 2, –3, 4 de L são os simétricos dos multiplicadores –2, 3, –4 das operações com as linhas utilizadas. (Para conferir, multiplique L por U e obtenha $A = LU$.)

3.42 Seja A a matriz do Problema 3.41. Encontre X_1, X_2, X_3, sendo X_i a solução de $AX = B_i$ nos casos (a) $B_1 = (1, 1, 1)$, (b) $B_2 = B_1 + X_1$, (c) $B_3 = B_2 + X_2$.

(a) Encontramos $L^{-1}B_i$ aplicando em B_1 as operações com as linhas (1), (2) e, depois, (3) dadas no Problema 3.41.

$$B_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \xrightarrow{(1)\text{ e }(2)} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \xrightarrow{(3)} \begin{bmatrix} 1 \\ -1 \\ 8 \end{bmatrix}$$

Substituindo para trás, resolvemos $UX = B$ com $B = (1, -1, 8)$ para obter $X_1 = (-25, 9, 8)$.

(b) Inicialmente, encontramos $B_2 = B_1 + X_1 = (1, 1, 1) + (-25, 9, 8) = (-24, 10, 9)$. Então, como antes,

$$B_2 = [-24, 10, 9]^T \xrightarrow{(1)\text{ e }(2)} [-24, 58, -63]^T \xrightarrow{(3)} [-24, 58, -295]^T$$

Substituindo para trás, resolvemos $UX = B$ com $B = (-24, 58, -295)$ para obter $X_2 = (943, -353, -295)$.

(c) Inicialmente, encontramos $B_3 = B_2 + X_2 = (-24, 10, 9) + (943, -353, -295) = (919, -343, -286)$. Então, como antes,

$$B_3 = [943, -353, -295]^T \xrightarrow{(1)\text{ e }(2)} [919, -2181, 2671]^T \xrightarrow{(3)} [919, -2181, 11\,395]^T$$

Substituindo para trás, resolvemos $UX = B$ com $B = (919, -2181, 11395)$ para obter

$$X_3 = (-37628, 13576, 11395).$$

Problemas variados

3.43 Seja L uma combinação linear das m equações com n incógnitas do sistema (3.2). Digamos que L seja a equação

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m \qquad (1)$$

Mostre que qualquer solução de (3.2) também é uma solução de L.
Seja $u = (k_1, ..., k_n)$ uma solução de (3.2). Então

$$a_{i1}k_1 + a_{i2}k_2 + \cdots + a_{in}k_n = b_i \qquad (i = 1, 2, \ldots, m) \qquad (2)$$

Substituindo u no lado esquerdo de (1) e usando (2), obtemos

$$\begin{aligned}(c_1a_{11} + \cdots + c_ma_{m1})k_1 \;+\; \cdots \;+\; (c_1a_{1n} + \cdots + c_ma_{mn})k_n \\ = c_1(a_{11}k_1 + \cdots + a_{1n}k_n) \;+\; \cdots \;+\; c_m(a_{m1}k_1 + \cdots + a_{mn}k_n) \\ = c_1b_1 + \cdots + c_mb_m\end{aligned}$$

Esse é o lado direito de (1), portanto, u é uma solução de (1).

3.44 Suponha que um sistema $\mathscr{M}$ de equações lineares tenha sido obtido de um sistema $\mathscr{L}$ por uma operação elementar (página 69). Mostre que $\mathscr{M}$ e $\mathscr{L}$ têm as mesmas soluções.

Cada equação L de $\mathscr{M}$ é uma combinação linear de equações de $\mathscr{L}$. Logo, pelo Problema 3.43, qualquer solução de $\mathscr{L}$ também é solução de $\mathscr{M}$. Por outro lado, cada operação elementar tem uma operação inversa, de modo que $\mathscr{L}$ pode ser obtido a partir de $\mathscr{M}$ por uma operação elementar. Isso significa que qualquer solução de $\mathscr{M}$ é uma solução de $\mathscr{L}$. Assim, $\mathscr{L}$ e $\mathscr{M}$ têm as mesmas soluções.

3.45 Demonstre o Teorema 3.4. Suponha que um sistema $\mathscr{M}$ de equações lineares seja obtido de um sistema $\mathscr{L}$ de equações lineares por uma sequência finita de operações elementares. Então $\mathscr{M}$ e $\mathscr{L}$ têm as mesmas soluções.

Cada passo da sequência mantém inalterado o conjunto solução (Problema 3.44). Assim, o sistema original $\mathscr{L}$ e o sistema final $\mathscr{M}$ (bem como qualquer sistema intermediário) têm as mesmas soluções.

3.46 Dizemos que um sistema de equações lineares $\mathscr{L}$ é *consistente* se nenhuma combinação linear de suas equações é uma equação degenerada L de termo constante não nulo. Mostre que $\mathscr{L}$ é consistente se, e só se, $\mathscr{L}$ for redutível à forma escalonada.

Suponha que $\mathscr{L}$ seja redutível à forma escalonada. Então $\mathscr{L}$ tem alguma solução, que também deve ser uma solução de cada combinação linear de suas equações. Assim, L, que não tem soluções, não pode ser uma combinação linear das equações de $\mathscr{L}$. Assim, $\mathscr{L}$ é consistente.
Por outro lado, suponha que $\mathscr{L}$ não seja redutível à forma escalonada. Então, no processo de redução, deve surgir alguma equação degenerada L com termo constante não nulo que é uma combinação linear das equações de $\mathscr{L}$. Assim, $\mathscr{L}$ não é consistente.

3.47 Suponha que u e v sejam vetores distintos. Mostre que os vetores $u + a(u - v)$ são distintos para escalares a distintos.

Suponha que $u + a_1(u - v) = u + a_2(u - v)$. Basta mostrar que $a_1 = a_2$. Temos

$$a_1(u - v) = a_2(u - v), \text{ portanto, } (a_1 - a_2)(u - v) = 0$$

Como u e v são distintos, $u - v \neq 0$. Assim, $a_1 - a_2 = 0$ e, portanto, $a_1 = a_2$.

3.48 Suponha que AB esteja definido. Demonstre as afirmações seguintes.

(a) Se A tiver uma linha nula, então AB tem uma linha nula.

(b) Se B tiver uma coluna nula, então AB tem uma coluna nula.

(a) Sejam R_i a linha nula de A e $C_1, C_2, ..., C_n$ as colunas de B. Então a enésima linha de AB é

$$(R_iC_1, R_iC_2, \ldots, R_iC_n) = (0, 0, 0, \ldots, 0)$$

(b) B^T tem uma linha nula, portanto, $B^TA^T = (AB)^T$ tem uma linha nula. Logo, AB tem uma coluna nula.

Problemas Complementares

Equações lineares, sistemas 2 × 2

3.49 Decida quais dos sistemas seguintes são lineares.

(a) $3x - 4y + 2yz = 8$, (b) $ex + 3y = \pi$, (c) $2x - 3y + kz = 4$

3.50 Resolva (a) $\pi x = 2$, (b) $3x + 2 = 5x + 7 - 2x$, (c) $6x + 2 - 4x = 5 + 2x - 3$

3.51 Resolva cada um dos sistemas seguintes.

(a) $\begin{aligned} 2x + 3y &= 1 \\ 5x + 7y &= 3 \end{aligned}$ (b) $\begin{aligned} 4x - 2y &= 5 \\ -6x + 3y &= 1 \end{aligned}$ (c) $\begin{aligned} 2x - 4 &= 3y \\ 5y - x &= 5 \end{aligned}$ (d) $\begin{aligned} 2x - 4y &= 10 \\ 3x - 6y &= 15 \end{aligned}$

3.52 Considere os sistemas seguintes nas incógnitas x e y.

(a) $\begin{aligned} x - ay &= 1 \\ ax - 4y &= b \end{aligned}$ (b) $\begin{aligned} ax + 3y &= 2 \\ 12x + ay &= b \end{aligned}$ (c) $\begin{aligned} x + ay &= 3 \\ 2x + 5y &= b \end{aligned}$

Para quais valores de a o sistema dado tem solução única e para quais pares de valores (a, b) o sistema tem mais do que uma solução?

Sistemas de equações lineares gerais

3.53 Resolva.

(a) $\begin{aligned} x + y + 2z &= 4 \\ 2x + 3y + 6z &= 10 \\ 3x + 6y + 10z &= 17 \end{aligned}$ (b) $\begin{aligned} x - 2y + 3z &= 2 \\ 2x - 3y + 8z &= 7 \\ 3x - 4y + 13z &= 8 \end{aligned}$ (c) $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 5x + 8y + 19z &= 11 \end{aligned}$

3.54 Resolva.

(a) $\begin{aligned} x - 2y &= 5 \\ 2x + 3y &= 3 \\ 3x + 2y &= 7 \end{aligned}$ (b) $\begin{aligned} x + 2y - 3z + 2t &= 2 \\ 2x + 5y - 8z + 6t &= 5 \\ 3x + 4y - 5z + 2t &= 4 \end{aligned}$ (c) $\begin{aligned} x + 2y + 4z - 5t &= 3 \\ 3x - y + 5z + 2t &= 4 \\ 5x - 4y + 6z + 9t &= 2 \end{aligned}$

3.55 Resolva.

(a) $\begin{aligned} 2x - y - 4z &= 2 \\ 4x - 2y - 6z &= 5 \\ 6x - 3y - 8z &= 8 \end{aligned}$ (b) $\begin{aligned} x + 2y - z + 3t &= 3 \\ 2x + 4y + 4z + 3t &= 9 \\ 3x + 6y - z + 8t &= 10 \end{aligned}$

3.56 Considere os sistemas seguintes nas incógnitas x, y e z.

(a) $\begin{aligned} x - 2y &= 1 \\ x - y + az &= 2 \\ ay + 9z &= b \end{aligned}$ (b) $\begin{aligned} x + 2y + 2z &= 1 \\ x + ay + 3z &= 3 \\ x + 11y + az &= b \end{aligned}$ (c) $\begin{aligned} x + y + az &= 1 \\ x + ay + z &= 4 \\ ax + y + z &= b \end{aligned}$

Para quais valores de a o sistema dado tem solução única e para quais pares de valores (a, b) o sistema tem mais do que uma solução? O valor de b não influi no sistema ter, ou não, solução única. Por quê?

Combinações lineares, sistemas homogêneos

3.57 Escreva v como uma combinação linear de u_1, u_2, u_3, com

(a) $v = (4, -9, 2)$, $u_1 = (1, 2, -1)$, $u_2 = (1, 4, 2)$, $u_3 = (1, -3, 2)$;

(b) $v = (1, 3, 2)$, $u_1 = (1, 2, 1)$, $u_2 = (2, 6, 5)$, $u_3 = (1, 7, 8)$;

(c) $v = (1, 4, 6)$, $u_1 = (1, 1, 2)$, $u_2 = (2, 3, 5)$, $u_3 = (3, 5, 8)$.

3.58 Sejam $u_1 = (1, 1, 2)$, $u_2 = (1, 3, -2)$, $u_3 = (4, -2, -1)$ de $\mathbf{R}^3$. Mostre que u_1, u_2, u_3 são ortogonais e escreva v como uma combinação linear de u_1, u_2, u_3 se (a) $v = (5, -5, 9)$, (b) $v = (1, -3, 3)$, (a) $v = (1, 1, 1)$. (*Sugestão*: use coeficientes de Fourier.)

3.59 Encontre a dimensão e uma base do conjunto solução W de cada um dos sistemas homogêneos seguintes.

(a) $\begin{aligned} x - y + 2z &= 0 \\ 2x + y + z &= 0 \\ 5x + y + 4z &= 0 \end{aligned}$ (b) $\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned}$ (c) $\begin{aligned} x + 2y + 3z + t &= 0 \\ 2x + 4y + 7z + 4t &= 0 \\ 3x + 6y + 10z + 5t &= 0 \end{aligned}$

3.60 Encontre a dimensão e uma base do conjunto solução W de cada um dos sistemas seguintes.

(a) $\begin{aligned} x_1 + 3x_2 + 2x_3 - x_4 - x_5 &= 0 \\ 2x_1 + 6x_2 + 5x_3 + x_4 - x_5 &= 0 \\ 5x_1 + 15x_2 + 12x_3 + x_4 - 3x_5 &= 0 \end{aligned}$ (b) $\begin{aligned} 2x_1 - 4x_2 + 3x_3 - x_4 + 2x_5 &= 0 \\ 3x_1 - 6x_2 + 5x_3 - 2x_4 + 4x_5 &= 0 \\ 5x_1 - 10x_2 + 7x_3 - 3x_4 + 18x_5 &= 0 \end{aligned}$

Matrizes escalonadas, forma canônica por linhas

3.61 Reduza cada uma das matrizes seguintes à forma escalonada e depois à forma canônica por linhas.

(a) $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 9 \\ 1 & 5 & 12 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 5 \\ 3 & 6 & 3 & -7 & 7 \end{bmatrix}$, (c) $\begin{bmatrix} 2 & 4 & 2 & -2 & 5 & 1 \\ 3 & 6 & 2 & 2 & 0 & 4 \\ 4 & 8 & 2 & 6 & -5 & 7 \end{bmatrix}$

3.62 Reduza cada uma das matrizes seguintes à forma escalonada e depois à forma canônica por linhas.

(a) $\begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 4 & 3 & 5 & 5 & 7 \\ 3 & 6 & 4 & 9 & 10 & 11 \\ 1 & 2 & 4 & 3 & 6 & 9 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 3 & 8 & 12 \\ 0 & 0 & 4 & 6 \\ 0 & 2 & 7 & 10 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & 3 & 1 & 3 \\ 2 & 8 & 5 & 10 \\ 1 & 7 & 7 & 11 \\ 3 & 11 & 7 & 15 \end{bmatrix}$

3.63 Usando somente entradas 0 e 1, apresente todas as possíveis matrizes 2×2 em forma canônica por linhas.

3.64 Usando somente entradas 0 e 1, encontre o número n de todas as possíveis matrizes 3×3 em forma canônica por linhas.

Matrizes elementares, aplicações

3.65 Sejam e_1, e_2, e_3, respectivamente, as operações elementares com as linhas seguintes.

"Trocar as linhas R_2 e R_3 entre si", "Substituir R_2 por $3R_2$" e "Substituir R_1 por $2R_3 + R_1$"

(a) Encontre as matrizes elementares E_1, E_2, E_3 correspondentes.

(b) Encontre as operações inversas e_1^{-1}, e_2^{-1}, e_3^{-1}; suas matrizes elementares E'_1, E'_2, E'_3; correspondentes e a relação entre essas matrizes e E_1, E_2, E_3.

(c) Descreva as operações elementares f_1, f_2, f_3 com as colunas.

(d) Encontre matrizes elementares F_1, F_2, F_3 correspondentes a f_1, f_2, f_3 e a relação entre essas matrizes e E_1, E_2, E_3.

3.66 Expresse cada uma das matrizes seguintes como um produto de matrizes elementares.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 6 \\ -3 & -7 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 3 & 8 & 7 \end{bmatrix}$$

3.67 Encontre a inversa de cada uma das matrizes seguintes (se existir).

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & -3 & 1 \\ 3 & -4 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 1 \\ 3 & 10 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{bmatrix}$$

3.68 Encontre a inversa de cada uma das matrizes $n \times n$ seguintes.

(a) A tem entradas 1 na diagonal e na *superdiagonal* (as entradas imediatamente acima da diagonal) e zeros nas demais entradas.

(b) B tem entradas 1 na diagonal e acima da diagonal e zeros nas demais entradas.

Fatoração LU

3.69 Encontre a fatoração LU das matrizes seguintes.

$$\text{(a)} \begin{bmatrix} 1 & -1 & -1 \\ 3 & -4 & -2 \\ 2 & -3 & -2 \end{bmatrix}, \text{(b)} \begin{bmatrix} 1 & 3 & -1 \\ 2 & 5 & 1 \\ 3 & 4 & 2 \end{bmatrix}, \text{(c)} \begin{bmatrix} 2 & 3 & 6 \\ 4 & 7 & 9 \\ 3 & 5 & 4 \end{bmatrix}, \text{(d)} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 10 \end{bmatrix}$$

3.70 Seja A a matriz do Problema 3.69(b). Encontre X_1, X_2, X_3, X_4, sendo

(a) X_1 é a solução de $AX = B_1$, com $B_1 = (1, 1, 1)^T$.

(b) Para $k > 1$, X_k é a solução de $AX = B_k$, com $B_k = B_{k-1} + X_{k-1}$.

3.71 Seja B a matriz do Problema 3.69(b) Encontre a fatoração LDU de B.

Problemas variados

3.72 Considere o sistema seguinte de incógnitas x e y.

$$\text{(a)} \quad \begin{aligned} ax + by &= 1 \\ cx + dy &= 0 \end{aligned} \qquad \text{(b)} \quad \begin{aligned} ax + by &= 0 \\ cx + dy &= 1 \end{aligned}$$

Suponha que $D = ad - bc \neq 0$. Mostre que cada sistema tem a solução única
(a) $x = d/D,\ y = -c/D$, (b) $x = -b/D,\ y = a/D$.

3.73 Encontre a inversa da operação com as linhas "Substituir R_i por $kR_j + k'R_i$ $(k' \neq 0)$".

3.74 Demonstre que suprimindo a última coluna de uma forma escalonada (respectivamente, da forma canônica) de uma matriz aumentada $M = [A, B]$ resulta um sistema em forma escalonada (respectivamente, em forma canônica) de A.

3.75 Sejam e uma operação elementar com as linhas e E sua matriz elementar associada, f a operação com as colunas correspondente e F sua matriz elementar. Demonstre que

(a) $f(A) = (e(A^T))^T$, (b) $F = E^T$, (c) $f(A) = AF$.

3.76 A matriz A é equivalente à matriz B, e escrevemos $A \approx B$, se existirem matrizes não singulares P e Q tais que $B = PAQ$. Demonstre que $\approx$ é uma relação de equivalência, ou seja, que valem

(a) $A \approx A$, (b) Se $A \approx B$, então $B \approx A$, (c) Se $A \approx B$ e $B \approx C$, então $A \approx C$.

Respostas dos Problemas Complementares

Notação $A = [R_1; R_2; \ldots]$ denota a matriz A de colunas $R_1, R_2, \ldots$ Os elementos de cada linha são separados por vírgulas (que podem ser omitidas no caso de números de um algarismo), as linhas são separadas por ponto e vírgula e 0 denota uma linha nula. Por exemplo,

$$A = [1, 2, 3, 4; \quad 5, -6, 7, -8; \quad 0] = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & -6 & 7 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

3.49 (a) não é, (b) é, (c) é linear em x, y, z, mas não em x, y, z, k.

3.50 (a) $x = 2/\pi$, (b) não tem solução (c) qualquer escalar é uma solução

3.51 (a) (2, –1), (b) não tem solução (c) (5, 2), (d) $(5 - 2a, a)$

3.52 (a) $a \neq \pm 2, \quad (2, 2), \quad (-2, -2)$, (b) $a \neq \pm 6, \quad (6, 4), \quad (-6, -4)$, (c) $a \neq \frac{5}{2}, \quad (\frac{5}{2}, 6)$

3.53 (a) $(2, 1, \frac{1}{2})$, (b) não tem solução, (c) $u = (-7a - 1, 2a + 2, a)$

3.54 (a) (3, –1), (b) $u = (-a + 2b, 1 + 2a - 2b, a, b)$, (c) não tem solução.

3.55 (a) $u = (\frac{1}{2}a + 2, \quad a, \quad \frac{1}{2})$, (b) $u = (\frac{1}{2}(7 - 5b - 4a), \quad a, \; \frac{1}{2}(1 + b), \; b)$

3.56 (a) $a \neq \pm 3, \quad (3, 3), \quad (-3, -3)$, (b) $a \neq 5$ e $a \neq -1, \quad (5, 7), \quad (-1, -5)$,
(c) $a \neq 1$ e $a \neq -2, \quad (-2, 5)$

3.57 (a) 2, –1, 3, (b) 6, –3, 1, (c) não é combinação linear

3.58 (a) 3, –2, 1, (b) $\frac{2}{3}, -1, \frac{1}{3}$, (c) $\frac{2}{3}, \frac{1}{7}, \frac{1}{21}$

3.59 (a) $\dim W = 1$, $u_1 = (-1, 1, 1)$, (b) $\dim W = 0$, não tem base,
(c) $\dim W = 2$, $u_1 = (-2, 1, 0, 0)$, $u_2 = (5, 0, -2, 1)$

3.60 (a) $\dim W = 3$, $u_1 = (-3, 1, 0, 0, 0)$, $u_2 = (7, 0, -3, 1, 0)$, $u_3 = (3, 0, -1, 0, 1)$,
(b) $\dim W = 2$, $u_1 = (2, 1, 0, 0, 0)$, $u_2 = (5, 0, -5, -3, 1)$

3.61 (a) $[1, 0, -\frac{1}{2}; \quad 0, 1, \frac{5}{2}; \quad 0]$, (b) [1, 2, 0, 0, 2; 0, 0, 1, 0, 5; 0, 0, 0, 1, 2],
(c) $[1, 2, 0, 4, -5, 3; \quad 0, 0, 1, -5, \frac{15}{2}, -\frac{5}{2}; \quad 0]$

3.62 (a) [1, 2, 0, 0, –4, –2; 0, 0, 1, 0, 1, 2; 0, 0, 0, 1, 2, 1; 0],
(b) [0, 1, 0, 0; 0, 0, 1, 0; 0, 0, 0, 1; 0], (c) [1, 0, 0, 4; 0, 1, 0, –1; 0, 0, 1, 2; 0]

3.63 5; [1, 0; 0, 1], [1, 1; 0, 0], [1, 0; 0, 0], [0, 1; 0, 0], 0

3.64 16

3.65 (a) [1, 0, 0; 0, 0, 1; 0, 1, 0], [1, 0, 0; 0, 3, 0; 0, 0, 1], [1, 0 2; 0, 1, 0; 0, 0, 1],
(b) $R_2 \leftrightarrow R_3; \quad \frac{1}{3}R_2 \to R_2; \quad -2R_3 + R_1 \to R_1$; cada $E_i' = E_i^{-1}$,
(c) $C_2 \leftrightarrow C_3, 3C_2 \to C_2, 2C_3 + C_1 \to C_1$, (d) cada $F_i = E_i^T$.

3.66 $A = [1, 0; \quad 3, 1][1, 0; \quad 0, -2][1, 2; \quad 0, 1]$, B não é invertível,
$C = [1, 0; \quad -\frac{3}{2}, 1][1, 0; \quad 0, 2][1, 6; \quad 0, 1][2, 0; \quad 0, 1]$,
$D = [100; \quad 010; \quad 301][100; \quad 010; \quad 021][100; \quad 013; \quad 001][120; \quad 010; \quad 001]$

3.67 $A^{-1} = [-8, 12, -5; \quad -5, 7, -3; \quad 1, -2, 1]$, B não tem inversa,
$C^{-1} = [\frac{29}{2}, -\frac{17}{2}, \frac{7}{2}; \quad -\frac{5}{2}, \frac{3}{2}, -\frac{1}{2}; \quad 3, -2, 1]$, $D^{-1} = [8, -3, -1; \quad -5, 2, 1; \quad 10, -4, -1]$

3.68 $A^{-1} = [1, -1, 1, -1, \ldots; \quad 0, 1, -1, 1, -1, \ldots; \quad 0, 0, 1, -1, 1, -1, 1, \ldots; \quad \ldots; \quad \ldots; \quad 0, \ldots 0, 1]$, B^{-1} tem entradas 1 na diagonal, –1 na superdiagonal e 0 no resto.

3.69 (a) $[100; \quad 310; \quad 211][1, -1, -1; \quad 0, -1, 1; \quad 0, 0, -1]$,
(b) $[100; \quad 210; \quad 351][1, 3, -1; \quad 0, -1, 3; \quad 0, 0, -10]$,
(c) $[100; \quad 210; \quad \frac{3}{2}, \frac{1}{2}, 1][2, 3, 6; \quad 0, 1, -3; \quad 0, 0, -\frac{7}{2}]$,
(d) Não há fatoração LU

3.70 $X_1 = [1, 1, -1]^T, \quad B_2 = [2, 2, 0]^T, \quad X_2 = [6, 4, 0]^T, \quad B_3 = [8, 6, 0]^T, \quad X_3 = [22, 16, -2]^T,$
$B_4 = [30, 22, -2]^T, \quad X_4 = [86, 62, -6]^T$

3.71 $B = [100; \quad 210; \quad 351] \operatorname{diag}(1, -1, -10) \, [1, 3, -1; \quad 0, 1, 3; \quad 0, 0, 1]$

3.73 Substituir R_i por $-kR_j + (1/k')R_i$.

3.75 (c) $f(A) = (e(A^T))^T = (EA^T)^T = (A^T)^T E^T = AF$

3.76 (a) $A = IAI$. (b) Se $A = PBQ$, então $B = P^{-1}AQ^{-1}$.
(c) Se $A = PBQ$ e $B = P'CQ'$, então $A = (PP')C(Q'Q)$.

Capítulo 4

Espaços Vetoriais

4.1 INTRODUÇÃO

Neste capítulo introduzimos a estrutura básica da Álgebra Linear, que é a do espaço vetorial de dimensão finita. A definição de um espaço vetorial V, cujos elementos são denominados *vetores*, envolve um corpo K arbitrário, cujos elementos são denominados *escalares*. Adotamos a notação seguinte (salvo menção em contrário).

V	o espaço vetorial dado
u, v, w	vetores de V
K	o corpo numérico dado
a, b, c ou k, r	escalares de K

Quase nada de essencial se perde considerando que K seja o corpo real **R** ou o corpo complexo **C**.

O leitor deve suspeitar que a reta real **R** tem "dimensão" um, que o plano cartesiano $\mathbf{R}^2$ tem "dimensão" dois e que o espaço $\mathbf{R}^3$ tem "dimensão" três. Neste capítulo formalizamos essa noção de "dimensão", que comprova a intuição do leitor.

Ao longo deste livro utilizamos a notação de conjuntos apresentada a seguir.

$a \in A$	O elemento a pertence ao conjunto A
$a, b \in A$	Os elementos a e b pertencem ao conjunto A
$\forall x \in A$	Para cada x de A
$\exists x \in A$	Existe algum x de A
$A \subseteq B$	A é um subconjunto de B
$A \cap B$	A interseção de A e B
$A \cup B$	A união de A e B
$\emptyset$	O conjunto vazio

4.2 ESPAÇOS VETORIAIS

A noção de espaço vetorial V sobre o corpo K de escalares é definida como segue.

DEFINIÇÃO Seja V um conjunto não vazio com as duas operações seguintes.

(i) **Adição de vetores** Associa a quaisquer $u, v \in V$ a *soma* $u + v$ em V.
(ii) **Multiplicação por escalar** Associa a quaisquer $u \in V,\ r \in K$ o *produto* $ru \in V$.

Então dizemos que V é um *espaço vetorial* (sobre o corpo K) se os axiomas seguintes forem verdadeiros.

[A$_1$] Para quaisquer vetores $u, v, w \in V$, $(u + v) + w = u + (v + w)$.
[A$_2$] Existe um vetor em V, que denotamos 0 e denominamos *vetor nulo*, ou *vetor zero*, tal que $u + 0 = u = 0 + u$, para qualquer vetor $u \in V$.
[A$_3$] Para cada vetor $u \in V$ existe um vetor em V, que denotamos $-u$ e denominamos *simétrico de u*, tal que

$$u + (-u) = 0 = (-u) + u.$$

[A$_4$] Para quaisquer vetores $u, v \in V$, $u + v = v + u$.
[M$_1$] Para quaisquer vetores $u, v \in V$ e escalar $r \in K$, $r(u + v) = rv + ru$.
[M$_2$] Para quaisquer escalares $a, b \in K$ e vetor $u \in V$, $(a + b)u = au + bu$.
[M$_3$] Para quaisquer escalares $a, b \in K$ e vetor $u \in V$, $(ab)u = a(bu)$.
[M$_4$] Para qualquer vetor $u \in V$, $1u = u$, onde 1 é o escalar unitário de K.

Esses oito axiomas se dividem naturalmente em duas partes (conforme indicamos nos rótulos). Os quatros primeiros dizem respeito somente à estrutura aditiva de V e podem ser resumidos dizendo que V é um *grupo comutativo* na adição. Isso significa que

(a) Qualquer soma finita $v_1 + v_2 + \cdots + v_m$ de vetores dispensa o uso de parênteses e independe da ordem das parcelas.
(b) O vetor nulo 0 é único, bem como o simétrico $-u$ de cada vetor u.
(c) (Lei do Cancelamento) Se $u + w = v + w$, então $u = v$.

Também definimos a *subtração* em V por $u - v = u + (-v)$, em que $-v$ é o único simétrico de v.

Por outro lado, os outros quatro axiomas dizem respeito à "ação" do corpo K de escalares sobre o espaço vetorial V. Usando esses axiomas, demonstramos (Problema 4.2) as propriedades simples de um espaço vetorial que seguem.

Teorema 4.1 Seja V um espaço vetorial sobre um corpo K.

(i) Para qualquer escalar $r \in K$ e $0 \in V$, $r0 = 0$.
(ii) Para $0 \in K$ e qualquer vetor $u \in V$, $0u = 0$.
(iii) Se $ru = 0$, com $r \in K$ e $u \in V$, então $r = 0$ ou $u = 0$.
(iv) Para quaisquer $r \in K$ e $u \in V$, $(-r)u = r(-u) = -ru$.

4.3 EXEMPLOS DE ESPAÇOS VETORIAIS

Nesta seção apresentamos exemplos de importantes espaços vetoriais que serão utilizados ao longo de todo o texto.

Espaço K^n

Seja K um corpo arbitrário. Frequentemente utilizamos a notação K^n para o conjunto de todas as ênuplas de elementos de K. Esse conjunto K^n é um espaço vetorial sobre K com as operações seguintes.

(i) **Adição de vetores** $(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1,\ a_2 + b_2, \ldots,\ a_n + b_n)$
(ii) **Multiplicação por escalar** $k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$

O vetor nulo de K^n é a ênupla de zeros,

$$0 = (0, 0, ..., 0)$$

e o simétrico de um vetor é definido por

$$-(a_1, a_2, \ldots, a_n) = (-a_1, -a_2, \ldots, -a_n)$$

Observe que essas são as mesmas operações já definidas para $\mathbf{R}^n$ no Capítulo 1. A demonstração de que K^n é um espaço vetorial é idêntica à demonstração do Teorema 1.1 e, agora, podemos afirmar que $\mathbf{R}^n$, com as operações lá definidas, é um espaço vetorial sobre $\mathbf{R}$.

Espaço polinomial P(*t*)

Seja **P**(t) o conjunto de todos os polinômios da forma

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s \qquad (s = 1, 2, \ldots)$$

em que os coeficientes a_i pertencem ao corpo K. Então **P**(t) é um espaço vetorial sobre K utilizando as operações seguintes.

(i) **Adição de vetores** A soma $p(t) + q(t)$ de **P**(t) é a operação de soma usual de polinômios.
(ii) **Multiplicação por escalar** O produto $kp(t)$ de **P**(t) é a operação de multiplicação usual de um escalar k e um polinômio $p(t)$.

O polinômio zero é o vetor nulo de **P**(t).

Espaço polinomial P$_n$(*t*)

Seja $\mathbf{P}_n(t)$ o conjunto de todos os polinômios $p(t)$ sobre um corpo K de grau menor do que ou igual a n, ou seja,

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s$$

com $s \leq n$. Então $\mathbf{P}_n(t)$ é um espaço vetorial sobre K em relação às operações usuais de soma de polinômios e de multiplicação de polinômios por uma constante (exatamente como no espaço vetorial **P**(t) precedente). Incluímos o polinômio zero 0 como um elemento de $\mathbf{P}_n(t)$, mesmo que não tenha grau algum.

Espaço matricial M$_{m,n}$

A notação $\mathbf{M}_{m,n}$ ou, simplesmente, **M**, é utilizada para denotar o conjunto de todas as matrizes $m \times n$ com entradas de um corpo K. Então $\mathbf{M}_{m,n}$ é um espaço vetorial sobre K em relação às operações usuais de soma matricial e de multiplicação de matrizes por escalares, conforme indicado no Teorema 2.1.

Espaço funcional *F* (*X*)

Sejam X um conjunto não vazio e K um corpo qualquer. O conjunto de todas as funções de X em K é denotado por $F(X)$. [Observe que $F(X)$ não é vazio, pois X não é vazio.] Então $F(X)$ é um espaço vetorial sobre K em relação às operações seguintes.

(i) **Adição de vetores** A soma de duas funções f e g de $F(X)$ é a função $f + g$ de $F(X)$ definida por

$$(f + g)(x) = f(x) + g(x) \qquad \forall x \in X$$

(ii) **Multiplicação por escalar** O produto de um escalar $k \in K$ e uma função f de $F(X)$ é a função kf de $F(X)$ definida por

$$(kf)(x) = kf(x) \qquad \forall x \in X$$

O vetor nulo de $F(X)$ é a função nula **0**, que leva cada $x \in X$ no elemento nulo $0 \in K$, ou seja,

$$\mathbf{0}(x) = 0 \qquad \forall x \in X$$

Também, dada qualquer função f de $F(X)$, a função simétrica de f é a função $-f$ de $F(X)$ definida por

$$(-f)(x) = -f(x) \qquad \forall x \in X$$

Corpos e subcorpos

Suponha que um corpo E seja uma extensão de um corpo K, ou seja, suponha que E seja um corpo que contém K como um subcorpo. Então E pode ser visto como um espaço vetorial sobre K com as operações seguintes.

(i) **Adição de vetores** A soma $u + v$ em E é a operação de soma usual de E.
(ii) **Multiplicação por escalar** O produto ru em E, com $r \in K$ e $u \in E$ é o produto usual de r e u como elementos de E.

Assim, os oito axiomas de um espaço vetorial são satisfeitos por E e seu subcorpo K em relação a essas duas operações.

4.4 COMBINAÇÕES LINEARES, CONJUNTOS GERADORES

Seja V um espaço vetorial sobre um corpo K. Um vetor v de V é uma *combinação linear* dos vetores $u_1, u_2, ..., u_m$ de V se existirem escalares $a_1, a_2, ..., a_m$ em K tais que

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

Alternativamente, v é uma combinação linear de $u_1, u_2, ..., u_m$ se existir alguma solução da equação vetorial

$$v = x_1u_1 + x_2u_2 + \cdots + x_mu_m$$

em que $x_1, x_2, ..., x_m$ são incógnitas escalares.

Exemplo 4.1 (Combinações Lineares em $\mathbf{R}^n$) Suponha que queiramos expressar $v = (3, 7, -4)$ de $\mathbf{R}^3$ como uma combinação linear dos vetores

$$u_1 = (1, 2, 3), \qquad u_2 = (2, 3, 7), \qquad u_3 = (3, 5, 6)$$

Procuramos escalares x, y, z tais que $v = xu_1 + yu_2 + zu_3$, ou seja,

$$\begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} + z\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix} \qquad \text{ou} \qquad \begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 5z &= 7 \\ 3x + 7y + 6z &= -4 \end{aligned}$$

(Por conveniência, escrevemos os vetores de $\mathbf{R}^3$ como colunas, porque dessa forma é mais fácil encontrar o sistema de equações lineares equivalente.) Reduzindo o sistema à forma escalonada, obtemos

$$\begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ y - 3z &= -13 \end{aligned} \qquad \text{e, então,} \qquad \begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ -4z &= -12 \end{aligned}$$

Substituindo para trás, obtemos a solução $x = 2$, $y = -4$, $z = 3$. Assim, $v = 2u_1 - 4u_2 + 3u_3$.

OBSERVAÇÃO De um modo genérico, a questão de expressar um dado vetor v de K^n como uma combinação linear de vetores $u_1, u_2, ..., u_m$ de K^n é equivalente a resolver um sistema $AX = B$ de equações lineares, onde v é a coluna B dos termos constantes e os vetores u_i são as colunas da matriz de coeficientes A. Um tal sistema pode ter uma única solução (como no exemplo), muitas soluções, ou não ter solução. Esse último caso, o de não ter solução, significa que v não pode ser escrito como combinação linear dos vetores $u_1, u_2, ..., u_m$.

Exemplo 4.2 (Combinações lineares em $\mathbf{P}(t)$) Suponha que queiramos expressar o polinômio $v = 3t^2 + 5t - 5$ como uma combinação linear dos polinômios

$$p_1 = t^2 + 2t + 1, \qquad p_2 = 2t^2 + 5t + 4, \qquad p_3 = t^2 + 3t + 6$$

Queremos encontrar escalares x, y, z tais que $v = xp_1 + yp_2 + zp_3$, ou seja,

$$3t^2 + 5t - 5 = x(t^2 + 2t + 1) + y(2t^2 + 5t + 4) + z(t^2 + 3t + 6) \qquad (*)$$

Há duas maneiras de proceder a partir daqui.

(1) Desenvolvemos o lado direito de (*) e obtemos

$$\begin{aligned} 3t^2 + 5t - 5 &= xt^2 + 2xt + x + 2yt^2 + 5yt + 4y + zt^2 + 3zt + 6z \\ &= (x + 2y + z)t^2 + (2x + 5y + 3z)t + (x + 4y + 6z) \end{aligned}$$

Igualamos os coeficientes de mesmas potências de t e reduzimos o sistema à forma escalonada.

$$\begin{aligned} x + 2y + z &= 3 \\ 2x + 5y + 3z &= 5 \\ x + 4y + 6z &= -5 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y + z &= 3 \\ y + z &= -1 \\ 2y + 5z &= -8 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y + z &= 3 \\ y + z &= -1 \\ 3z &= -6 \end{aligned}$$

O sistema está em forma triangular e tem uma solução. Substituindo para trás, obtemos a solução $x = 3$, $y = 1$, $z = -2$. Assim,

$$v = 3p_1 + p_2 - 2p_3$$

(2) Observamos que a equação (*) é, na verdade, uma identidade na variável t, ou seja, a equação é válida para qualquer valor de t. Podemos obter três equações nas incógnitas x, y, z igualando t a três valores quaisquer. Por exemplo,

tomando $t = 0$ em (*), obtemos $\quad x + 4y + 6z = -5$

tomando $t = 1$ em (*), obtemos $\quad 4x + 11y + 10z = 3$

tomando $t = -1$ em (*), obtemos $\quad y + 4z = -7$

Reduzindo esse sistema à forma escalonada e substituindo para trás, novamente obtemos a solução $x = 3$, $y = 1$, $z = -2$. Assim (de novo), $v = 3p_1 + p_2 - 2p_3$.

Conjuntos geradores

Seja V um espaço vetorial sobre K. Dizemos que os vetores $u_1, u_2, \ldots, u_m$ de V *geram* V, ou que constituem um *conjunto gerador* de V, se cada vetor v de V for uma combinação linear dos vetores $u_1, u_2, \ldots, u_m$, ou seja, se existirem escalares vetores $a_1, a_2, \ldots, a_m$ de K tais que

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

As observações seguintes decorrem diretamente da definição.

OBSERVAÇÃO 1 Suponha que $u_1, u_2, \ldots, u_m$ gerem V. Então, dado qualquer vetor w, os vetores $w, u_1, u_2, \ldots, u_m$ também geram V.

OBSERVAÇÃO 2 Suponha que $u_1, u_2, \ldots, u_m$ gerem V e que u_k seja uma combinação linear de alguns dos demais vetores u. Então esses outros u, sem o vetor u_k, também geram V.

OBSERVAÇÃO 3 Suponha que $u_1, u_2, \ldots, u_m$ gerem V e que um dos u seja o vetor nulo. Então os demais vetores u também geram V.

Exemplo 4.3 Considere o espaço vetorial $V = \mathbf{R}^3$

(a) Afirmamos que os vetores seguintes formam um conjunto gerador de $\mathbf{R}^3$.

$$e_1 = (1, 0, 0), \qquad e_2 = (0, 1, 0), \qquad e_3 = (0, 0, 1)$$

Mais precisamente, se $v = (a, b, c)$ é um vetor qualquer de $\mathbf{R}^3$, então

$$v = ae_1 + be_2 + ce_3$$

(b) Afirmamos que os vetores seguintes também formam um conjunto gerador de $\mathbf{R}^3$.

$$w_1 = (1, 1, 1), \qquad w_2 = (1, 1, 0), \qquad w_3 = (1, 0, 0)$$

Mais precisamente, se $v = (a, b, c)$ é um vetor qualquer de $\mathbf{R}^3$, então (Problema 4.62)

$$v = (a, b, c) = cw_1 + (b - c)w_2 + (a - b)w_3$$

Por exemplo, $v = (5, -6, 2) = 2w_1 - 8w_2 + 11w_3$.

(c) Podemos mostrar [Problema 3.24(b)] que $v = (2, 7, 10)$ não pode ser escrito como uma combinação linear dos vetores

$$u_1 = (1, 2, 3), \qquad u_2 = (1, 3, 5), \qquad u_3 = (1, 5, 9)$$

Por isso, u_1, u_2, u_3 não geram $\mathbf{R}^3$.

Exemplo 4.4 Considere o espaço vetorial $V = \mathbf{P}_n(t)$, que consiste em todos os polinômios de grau $\leq n$.

(a) Claramente, cada polinômio de $\mathbf{P}_n(t)$ pode ser expresso como uma combinação linear dos $n + 1$ polinômios

$$1, t, t^2, t^3, ..., t^n$$

Assim, essas potências de t (sendo $1 = t^0$) formam um conjunto gerador de $\mathbf{P}_n(t)$.

(b) Também podemos mostrar que, dado qualquer escalar c, as $n + 1$ potências seguintes de $t - c$,

$$1, t - c, (t - c)^2, (t - c)^3, ..., (t - c)^n$$

(em que $(t - c)^0 = 1$), também formam um conjunto gerador de $\mathbf{P}_n(t)$.

Exemplo 4.5 Considere o espaço vetorial $\mathbf{M} = \mathbf{M}_{2,2}$, que consiste em todas as matrizes 2×2, e as quatro matrizes de $\mathbf{M}$ a seguir.

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Claramente, cada matriz A de $\mathbf{M}$ pode ser escrita como uma combinação linear dessas quatro matrizes. Por exemplo,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Por isso, as quatro matrizes $E_{11}, E_{12}, E_{21}, E_{22}$ geram $\mathbf{M}$.

4.5 SUBESPAÇOS

Nesta seção introduzimos a importante noção de subespaço vetorial.

DEFINIÇÃO Sejam V um espaço vetorial sobre o corpo K e W um subconjunto de V. Então W é um *subespaço* de V se o próprio W for um espaço vetorial sobre o corpo K em relação às operações de V de soma de vetores e de multiplicação de vetores por escalar.

Para mostrar que um conjunto W qualquer é um espaço vetorial, precisamos mostrar que W satisfaz os oito axiomas de um espaço vetorial. No entanto, se W for um subconjunto de um espaço vetorial V, então alguns desses axiomas valem automaticamente para W, pois já valem para V. Vejamos alguns critérios simples para identificar subespaços.

Teorema 4.2 Seja W um subconjunto de um espaço vetorial V. Então W é um subespaço de V se as duas condições seguintes são verdadeiras.

(a) O vetor nulo 0 pertence a W.
(b) Dados quaisquer $u, v \in W$ e $r \in K$, (i) a soma $u + v \in W$ e (ii) o múltiplo $ru \in W$.

A propriedade (i) de (b) diz que W é *fechado em relação à adição*, e a propriedade (ii) de (b) diz que W é *fechado em relação à multiplicação por escalar*. Essas duas propriedades podem ser combinadas numa única afirmação equivalente, como segue.

(b') Para quaisquer $u, v \in W, a, b \in K$, a combinação linear $au + bv \in W$.

Seja, agora, V um espaço vetorial qualquer. Então automaticamente V contém dois subespaços, o conjunto $\{0\}$ consistindo apenas no vetor nulo e o espaço V todo. Esses são, às vezes, denominados subespaços *triviais* de V. A seguir, alguns exemplos de subespaços não triviais.

Exemplo 4.6 Considere o espaço vetorial $V = \mathbf{R}^3$.

(a) Seja U o conjunto de todos os vetores de $\mathbf{R}^3$ cujas entradas são iguais, ou seja,

$$U = \{(a, b, c) : a = b = c\}$$

Por exemplo, (1, 1, 1), (–3, –3, –3), (7, 7, 7) e (–2, –2, –2) são vetores de U. Geometricamente, U é a reta pela origem O e pelo ponto (1, 1, 1), conforme Figura 4-1(a). Claramente, $0 = (0, 0, 0)$ pertence a U, porque todas as entradas de 0 são iguais. Além disso, sejam u e v dois vetores arbitrários de U, digamos, $u = (a, a, a)$ e $v = (b, b, b)$. Então, dado qualquer escalar $r \in K$, também são vetores de U os vetores a seguir.

$$u + v = (a + b, a + b, a + b) \qquad \text{e} \qquad ru = (ra, ra, ra)$$

Assim, U é um subespaço de $\mathbf{R}^3$.

(b) Seja W um plano pela origem arbitrário de $\mathbf{R}^3$, conforme a Figura 4-1(b). Então $0 = (0, 0, 0)$ pertence a W, porque estamos supondo que W passa pela origem O. Além disso, sejam u e v dois vetores de W. Então u e v podem ser interpretados como setas do plano W de ponto inicial na origem O, conforme a Figura 4-1(b). A soma $u + v$ e qualquer múltiplo ru de u também são setas do plano W. Assim, W é um subespaço de $\mathbf{R}^3$.

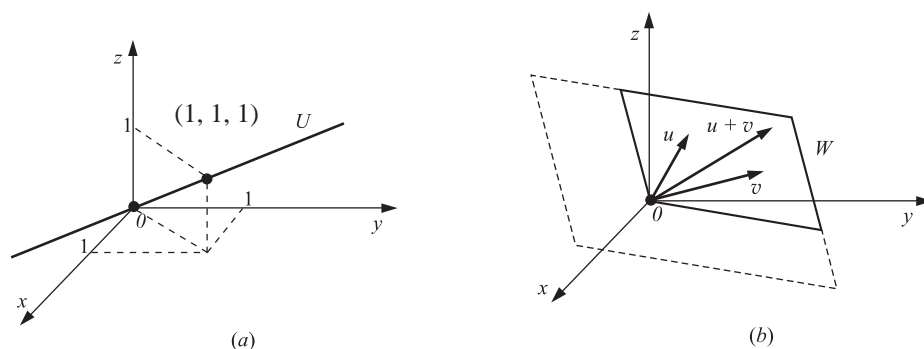



Figura 4-1

Exemplo 4.7

(a) Seja $V = \mathbf{M}_{n,n}$ o espaço vetorial das matrizes $n \times n$. Sejam W_1 o subconjunto de todas as matrizes triangulares (superiores) e W_2 o subconjunto de todas as matrizes simétricas. Então W_1 é um subespaço de V, porque W_1 contém a matriz nula 0 e W_1 é fechado em relação à soma de matrizes e à multiplicação por escalar, ou seja, a soma e a multiplicação por escalar de tais matrizes triangulares também é triangular. Analogamente, W_2 é um subespaço de V.

(b) Seja $V = \mathbf{P}(t)$, o espaço vetorial dos polinômios. Então o espaço $\mathbf{P}_n(t)$ dos polinômios de grau no máximo n pode ser visto como um subespaço de $\mathbf{P}(t)$. Seja $\mathbf{Q}(t)$ o conjunto de todos os polinômios que possuem apenas potências pares de t. Por exemplo, os polinômios seguintes são de $\mathbf{Q}(t)$.

$$p_1 = 3 + 4t^2 - 5t^6 \qquad \text{e} \qquad p_2 = 6 - 7t^4 + 9t^6 + 3t^{12}$$

(Estamos supondo que qualquer constante $k = kt^0$ é uma potência par de t.) Então $\mathbf{Q}(t)$ é um subespaço de $\mathbf{P}(t)$.

(c) Seja V o espaço vetorial das funções reais. Então o conjunto W_1 das funções contínuas e o conjunto W_2 das funções deriváveis são subespaços de V.

Interseção de subespaços

Sejam U e W subespaços de um espaço vetorial V. Mostremos que a interseção $U \cap W$ também é um subespaço de V. Claramente, $0 \in U$ e $0 \in W$, pois U e W são subespaços; logo, $0 \in U \cap W$. Agora suponha que u e v pertençam à interseção $U \cap W$. Então $u, v \in U$ e $u, v \in W$. Além disso, como U e W são subespaços, dados quaisquer escalares $a, b \in K$,

$$au + bv \in U \qquad \text{e} \qquad au + bv \in W$$

Assim, $au + bv \in U \cap W$. Logo, $U \cap W$ é um subespaço de V.

O resultado precedente é generalizado como segue.

Teorema 4.3 A interseção de um número qualquer de subespaços de um espaço vetorial V é um subespaço de V.

Espaço solução de um sistema homogêneo

Considere um sistema $AX = B$ de equações lineares com n incógnitas. Então cada solução u pode ser vista como um vetor de K^n. Assim, o conjunto solução de um tal sistema é um subconjunto de K^n. Suponha, agora, que o sistema seja homogêneo, isto é, da forma $AX = 0$. Seja W o conjunto solução dessa equação. Como $A0 = 0$, o vetor nulo $0 \in W$. Além disso, suponha que u e v pertençam a W. Então u e v são soluções de $AX = 0$, ou seja, $Au = 0$ e $Av = 0$. Portanto, dados quaisquer escalares a e b obtemos

$$A(au + bv) = aAu + bAv = a0 + b0 = 0 + 0 = 0$$

Assim, $au + bv$ pertence a W, por ser uma solução de $AX = 0$. Decorre disso que W é um subespaço de K^n.

Formalizamos o resultado precedente como segue.

Teorema 4.4 O conjunto solução W de um sistema homogêneo $AX = 0$ com n incógnitas é um subespaço de K^n.

Enfatizamos que o conjunto solução de um sistema não homogêneo $AX = B$ não é um subespaço de K^n. Na verdade, o vetor nulo 0 sequer pertence ao conjunto solução dessa equação.

4.6 ESPAÇOS GERADOS, ESPAÇO LINHA DE UMA MATRIZ

Suponha que $u_1, u_2, ..., u_m$ sejam vetores quaisquer de um espaço vetorial V. Na Seção 4.4 vimos que qualquer vetor da forma $a_1u_1 + a_2u_2 + \cdots + a_mu_m$, em que os a_i são escalares, é denominado uma *combinação linear* de $u_1, u_2, ..., u_m$. O conjunto de todas essas combinações lineares, denotado por

$$\text{ger}(u_1, u_2, ..., u_m) \qquad \text{ou} \qquad \text{ger}(u_i)$$

é denominado *espaço gerado* por $u_1, u_2, ..., u_m$.

Claramente, o vetor nulo 0 pertence ao ger(u_i), pois

$$0 = 0u_1 + 0u_2 + \cdots + 0u_m$$

Além disso, suponha que v e v' pertençam a ger(u_i), digamos,

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m \qquad \text{e} \qquad v' = b_1u_1 + b_2u_2 + \cdots + b_mu_m$$

Então

$$v + v' = (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_m + b_m)u_m$$

e, dado qualquer escalar $k \in K$,

$$kv = ka_1u_1 + ka_2u_2 + \cdots + ka_mu_m$$

Assim, $v + v'$ e kv também pertencem a ger(u_i). Decorre disso que ger(u_i) é um subespaço de V.

Mais geralmente, dado qualquer subconjunto S de V, ger(S) consiste em todas as combinações lineares de vetores de S ou, se $S = \phi$, ger$(S) = \{0\}$. Assim, em particular, S é um conjunto gerador de ger(S) (ver Seção 4.4).

Vale o teorema seguinte, que acabamos de demonstrar parcialmente.

Teorema 4.5 Seja S um subconjunto de um espaço vetorial V.

(i) Então ger(S) é um subespaço de V que contém S.
(ii) Se W é um subespaço de V que contém S, então $(S) \subseteq W$.

A condição (ii) do Teorema 4.5 pode ser interpretada como dizendo que ger(S) é o "menor" subespaço de V que contém S.

Exemplo 4.8 Considere o espaço vetorial $V = \mathbf{R}^3$.

(a) Seja u um vetor não nulo qualquer de $\mathbf{R}^3$. Então ger(u) consiste em todos os múltiplos escalares de u. Geometricamente, ger(u) é a reta pela origem O e pela extremidade de u, conforme Figura 4-2(a).

(a)

(b)

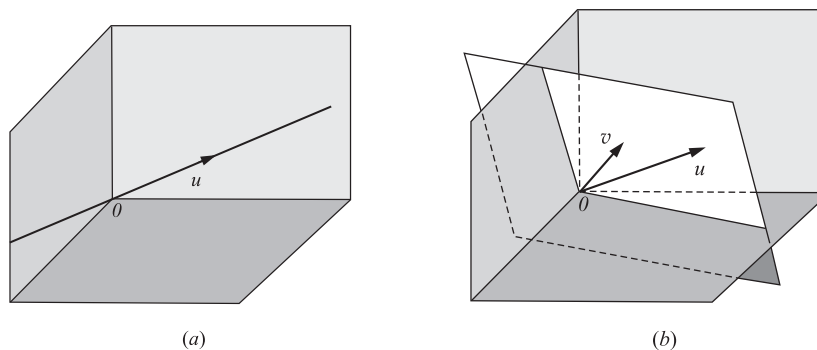



Figura 4-2

(b) Sejam u e v vetores de $\mathbf{R}^3$ que não são múltiplos um do outro. Então ger(u, v) é o plano pela origem O e pelas extremidades de u e v, conforme Figura 4-2(b).

(c) Considere os vetores $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ de $\mathbf{R}^3$. No Exemplo 4.3(a) vimos que cada vetor de $\mathbf{R}^3$ é uma combinação linear de e_1, e_2, e_3. Ou seja, e_1, e_2, e_3 é um conjunto gerador de $\mathbf{R}^3$. Decorre disso que ger$(e_1, e_2, e_3) = \mathbf{R}^3$.

Espaço linha de uma matriz

Seja $A = [a_{ij}]$ uma matriz $m \times n$ qualquer sobre um corpo K. As linhas de A,

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \qquad R_2 = (a_{21}, a_{22}, \ldots, a_{2n}), \qquad \ldots, \qquad R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

podem ser vistas como vetores de K^n e, portanto, geram um subespaço de K^n, denominado *espaço linha* de A e denotado por lin(A). Temos

$$\text{lin}(A) = \text{ger}(R_1, R_2, \ldots, R_m)$$

Analogamente, as colunas de A podem ser vistas como vetores de K^m que geram um subespaço de K^m denominado *espaço coluna* de A e denotado por col(A). Observe que col(A) = lin(A^T).

Lembre-se que as matrizes A e B são ditas equivalentes por linhas, e escrevemos $A \sim B$, se B pode ser obtida a partir de A por uma sequência de operações elementares com as linhas. Agora suponha que M seja a matriz obtida aplicando uma das operações elementares seguintes a A.

(1) Trocar R_i e R_j entre si (2) Substituir R_i por kR_i (3) Substituir R_j por $kR_i + R_j$

Então cada linha de M ou é uma linha de A ou é uma combinação linear de linhas de A. Assim, o espaço linha de M está contido no espaço linha de A. Por outro lado, podemos aplicar as operações elementares inversas a M para obter A; portanto, o espaço linha de A está contido no espaço linha de M. Isso mostra que A e M têm o mesmo espaço linha. Isso vale cada vez que aplicarmos uma operação elementar com as linhas. Assim, demonstramos o teorema seguinte.

Teorema 4.6 Matrizes equivalentes por linhas têm o mesmo espaço linha.

Agora estamos em condições de demonstrar (Problemas 4.45–4.47) resultados básicos relativos à equivalência por linhas (que apareceram pela primeira vez nos Teoremas 3.7 e 3.8 do Capítulo 3).

Teorema 4.7 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ duas matrizes escalonadas equivalentes por linhas com respectivas entradas pivô

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r} \qquad \text{e} \qquad b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$$

Então A e B têm o mesmo número de linhas não nulas, ou seja, $r = s$, e as entradas pivô estão nas mesmas posições, ou seja, $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$.

Teorema 4.8 Sejam A e B matrizes em forma canônica por linhas. Então A e B têm o mesmo espaço linha se, e só se, A e B têm as mesmas linhas não nulas.

Corolário 4.9 Toda matriz A é equivalente por linhas a uma única matriz em forma canônica por linhas.

No próximo exemplo utilizamos esses resultados.

Exemplo 4.9 Considere os dois conjuntos de vetores de $\mathbf{R}^4$ dados a seguir.

$$u_1 = (1, 2, -1, 3), \qquad u_2 = (2, 4, 1, -2), \qquad u_3 = (3, 6, 3, -7)$$
$$w_1 = (1, 2, -4, 11), \qquad w_2 = (2, 4, -5, 14)$$

Sejam $U = \text{ger}(u_i)$ e $W = \text{ger}(w_i)$. Há duas maneiras de mostrar que $U = W$.

(a) Mostramos que cada u_i é uma combinação linear de w_1 e w_2 e, depois, que cada w_i é uma combinação linear de u_1, u_2, u_3. Observe que, para isso, devemos mostrar que são consistentes seis sistemas de equações lineares.

(b) Formamos a matriz A cujas linhas são u_1, u_2, u_3 e reduzimos A à forma canônica por linhas, formamos a matriz B cujas linhas são w_1 e w_2 e reduzimos B à forma canônica por linhas, como segue.

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{bmatrix}$$

Como as linhas não nulas dessas matrizes em forma canônica por linhas são idênticas, os espaços linha de A e de B são iguais. Portanto, $U = W$.

Claramente, o método de (b) é mais eficiente do que o método de (a).

4.7 DEPENDÊNCIA E INDEPENDÊNCIA LINEAR

Seja V um espaço vetorial sobre um corpo K. As noções de dependência e independência linear de vetores de V sobre K são definidas como segue. (Geralmente omitimos a menção a K quando o corpo está subentendido.) Esse conceito desempenha um papel essencial na teoria da Álgebra Linear e na Matemática em geral.

DEFINIÇÃO Dizemos que os vetores $v_1, v_2, \ldots, v_m$ de V são *linearmente dependentes* se existem escalares $a_1, a_2, \ldots, a_m$ de K, não todos nulos, tais que

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0$$

Caso contrário, dizemos que os vetores são *linearmente independentes*.

Podemos reformular essa definição como segue. Considere a equação vetorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m = 0 \qquad (*)$$

em que os x são incógnitas escalares. Essa equação sempre tem a solução nula $x_1 = 0, x_2 = 0, \ldots, x_m = 0$. Suponha que essa seja a única solução, isto é, suponha que possamos mostrar que

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m = 0 \quad \text{implica} \quad x_1 = 0, x_2 = 0, \ldots, x_m = 0$$

Então os vetores $v_1, v_2, \ldots, v_m$ são linearmente independentes. Por outro lado, suponha que a equação (*) possua alguma solução não nula; então os vetores são linearmente dependentes.

Um conjunto $S = \{v_1, v_2, \ldots, v_m\}$ de vetores de V é dito linearmente dependente ou independente se os vetores $v_1, v_2, \ldots, v_m$ são linearmente dependentes ou independentes.

Um conjunto infinito S de vetores é linearmente dependente ou independente se existirem, ou não, vetores $v_1, v_2, \ldots, v_k$ de S que são linearmente dependentes.

Atenção Um conjunto $S = \{v_1, v_2, ..., v_m\}$ como dado acima representa uma *lista* ou, em outras palavras, uma sequência finita de vetores que são ordenados e que podem ser repetidos.

As observações seguintes decorrem diretamente da definição dada.

OBSERVAÇÃO 1 Suponha que 0 seja um dos vetores $v_1, v_2, ..., v_m$, digamos, $v_1 = 0$. Então os vetores são necessariamente linearmente dependentes, porque temos a combinação linear seguinte em que o coeficiente de v_1 é não nulo.

$$1v_1 + 0v_2 + \cdots + 0v_m = 1 \cdot 0 + 0 + \cdots + 0 = 0$$

OBSERVAÇÃO 2 Suponha que v seja um vetor não nulo. Então v, sozinho, é linearmente independente porque

$$kv = 0, \qquad v \neq 0 \qquad \text{implica} \qquad k = 0$$

OBSERVAÇÃO 3 Suponha que dois dos vetores $v_1, v_2, ..., v_m$ sejam iguais, ou que um deles é múltiplo de um outro, digamos, $v_1 = kv_2$. Então os vetores são necessariamente linearmente dependentes, porque temos a combinação linear seguinte em que o coeficiente de v_1 é não nulo.

$$v_1 - kv_2 + 0v_3 + \cdots + 0v_m = 0$$

OBSERVAÇÃO 4 Dois vetores v_1 e v_2 são linearmente dependentes se, e só se, um deles é um múltiplo do outro.

OBSERVAÇÃO 5 Se o conjunto $\{v_1, v_2, ..., v_m\}$ for linearmente independente, então qualquer reordenação $\{v_{i_1}, v_{i_2}, \ldots, v_{i_m}\}$ de seus vetores ainda será linearmente independente.

OBSERVAÇÃO 6 Se um conjunto S de vetores for linearmente independente, então qualquer subconjunto de S é linearmente independente. Alternativamente, se S contiver um subconjunto linearmente dependente, então S é linearmente dependente.

Exemplo 4.10

(a) Sejam $u = (1, 1, 0)$, $v = (1, 3, 2)$, $w = (4, 9, 5)$. Então u, v, w são linearmente dependentes, pois

$$3u + 5v - 2w = 3(1, 1, 0) + 5(1, 3, 2) - 2(4, 9, 5) = (0, 0, 0) = 0$$

(b) Mostramos que os vetores $u = (1, 2, 3)$, $v = (2, 5, 7)$, $w = (1, 3, 5)$ são linearmente independentes. Para isso, formamos a equação vetorial $xu + yv + zw = 0$, em que x, y, z são incógnitas escalares. Isso fornece

$$x\begin{bmatrix}1\\2\\3\end{bmatrix} + y\begin{bmatrix}2\\5\\7\end{bmatrix} + z\begin{bmatrix}1\\3\\5\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 0 \\ 2x + 5y + 3z &= 0 \\ 3x + 7y + 5z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 0 \\ y + z &= 0 \\ 2z &= 0 \end{aligned}$$

Substituindo para trás, obtemos $x = 0, y = 0, z = 0$. Mostramos que

$$xu + yv + zw = 0 \quad \text{implica} \quad x = 0, y = 0, z = 0$$

Assim, u, v, w são linearmente independentes.

(c) Seja V o espaço vetorial das funções de **R** em **R**. Mostremos que as funções $f(t) = \operatorname{sen} t$, $g(t) = e^t$ e $h(t) = t^2$ são linearmente independentes. Para isso, formamos a equação vetorial funcional $xf + yg + zh = 0$, em que x, y, z são incógnitas escalares. Essa equação funcional significa que, para cada valor de t,

$$x \operatorname{sen} t + ye^t + zt^2 = 0$$

Assim, nessa equação, escolhemos valores apropriados de t para obter $x = 0, y = 0, z = 0$ facilmente. Por exemplo,

(i)	Substituindo $t = 0$	obtemos $x(0) + y(1) + z(0) = 0$	ou	$y = 0$
(ii)	Substituindo $t = \pi$	obtemos $x(0) + 0(e^{\pi}) + z(\pi^2) = 0$	ou	$z = 0$
(iii)	Substituindo $t = \pi/2$	obtemos $x(1) + 0(e^{\pi/2}) + 0(\pi^2/4) = 0$	ou	$x = 0$

Mostramos que

$$xf + yg + zh = 0 \qquad \text{implica} \qquad x = 0, y = 0, z = 0$$

Assim, f, g, h são linearmente independentes.

Dependência linear em $\mathbf{R}^3$

A dependência linear no espaço vetorial $V = \mathbf{R}^3$ pode ser descrita geometricamente como segue.

(a) Dois vetores quaisquer u e v de $\mathbf{R}^3$ são linearmente dependentes se, e só se, estão na mesma reta pela origem O, conforme Figura 4-3(a).
(b) Três vetores quaisquer u, v, w de $\mathbf{R}^3$ são linearmente dependentes se, e só se, estão no mesmo plano pela origem O, conforme Figura 4-3(b).

Adiante poderemos mostrar que quatro ou mais vetores quaisquer de $\mathbf{R}^3$ são automaticamente linearmente dependentes.

(a) u e v são linearmente dependentes

(b) u, v, w são linearmente dependentes

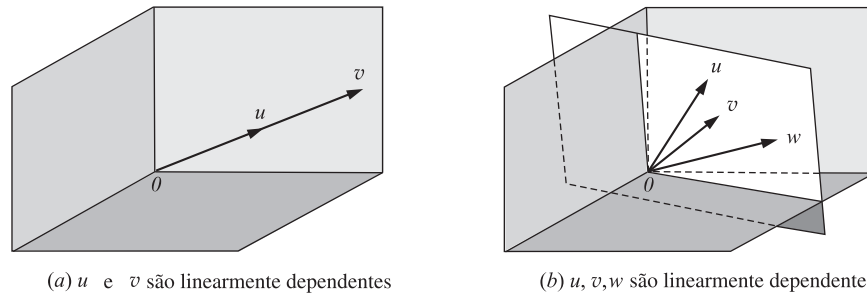



Figura 4-3

Dependência linear e combinações lineares

As noções de dependência linear e combinações lineares estão relacionadas de perto. Mais precisamente, para mais do que um vetor, mostramos que os vetores $v_1, v_2, ..., v_m$ são linearmente dependentes se, e só se, algum deles é uma combinação linear dos demais.

Suponha, por exemplo, que v_i seja uma combinação linear dos demais,

$$v_i = a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_m v_m$$

Então, somando $-v_i$ a ambos lados, obtemos

$$a_1 v_1 + \cdots + a_{i-1} v_{i-1} - v_i + a_{i+1} v_{i+1} + \cdots + a_m v_m = 0$$

com o coeficiente de v_i não sendo nulo. Portanto, os vetores são linearmente dependentes. Reciprocamente, suponha que os vetores sejam linearmente dependentes, digamos,

$$b_1 v_1 + \cdots + b_j v_j + \cdots + b_m v_m = 0, \qquad \text{com} \qquad b_j \neq 0$$

Então podemos resolver em v_j e obter

$$v_j = b_j^{-1} b_1 v_1 - \cdots - b_j^{-1} b_{j-1} v_{j-1} - b_j^{-1} b_{j+1} v_{j+1} - \cdots - b_j^{-1} b_m v_m$$

de modo que v_j é uma combinação linear dos outros vetores.

A seguir enunciamos uma afirmação um pouco mais forte do que a que acabamos de provar (Problema 4.33). Esse resultado tem muitas consequências importantes.

Lema 4.10 Suponha que dois ou mais vetores $v_1, v_2, ..., v_m$ sejam linearmente dependentes. Então algum dos vetores é uma combinação linear dos vetores que o antecedem, ou seja, existe algum $k > 1$ tal que

$$v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1}$$

Dependência linear e matrizes escalonadas

Considere a seguinte matriz escalonada A, cujos pivôs aparecem circulados.

$$A = \begin{bmatrix} 0 & \textcircled{2} & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & \textcircled{4} & 3 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & \textcircled{7} & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{6} & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe que as linhas R_2, R_3, R_4 têm zeros na segunda coluna abaixo do pivô não nulo de R_1 e, portanto, qualquer combinação linear de R_2, R_3, R_4 necessariamente tem 0 como segunda entrada. Assim, R_1 não pode ser uma combinação linear das linhas abaixo dela. Analogamente, as linhas R_3 e R_4 têm zeros na terceira coluna abaixo do pivô não nulo de R_2 e, portanto, R_2 não pode ser uma combinação linear das linhas abaixo dela. Finalmente, R_3 não pode ser um múltiplo de R_4, porque R_4 tem um 0 na quinta coluna abaixo do pivô não nulo de R_3. Considerando as linhas de baixo para cima, R_4, R_3, R_2, R_1, nenhuma linha é uma combinação linear das linhas que a precedem nessa lista. Assim, as linhas são linearmente independentes, pelo Lema 4.10.

O argumento utilizado com essa matriz escalonada pode ser utilizado para as linhas não nulas de qualquer matriz escalonada. Assim, obtemos o seguinte resultado útil.

Teorema 4.11 As linhas não nulas de uma matriz escalonada são linearmente independentes.

4.8 BASE E DIMENSÃO

Inicialmente enunciamos duas maneiras equivalentes de definir uma base de um espaço vetorial V. (A equivalência está demonstrada no Problema 4.28.)

DEFINIÇÃO A Um conjunto $S = \{u_1, u_2, ..., u_n\}$ de vetores é uma *base* de V se tiver as duas propriedades seguintes: (i) S é linearmente independente. (ii) S gera V.

DEFINIÇÃO B Um conjunto $S = \{u_1, u_2, ..., u_n\}$ de vetores é uma *base* de V se cada $v \in V$ puder ser escrito de maneira única como uma combinação linear dos vetores da base.

O seguinte é um resultado fundamental da Álgebra Linear.

Teorema 4.12 Seja V um espaço vetorial tal que alguma base tem m elementos e outra base tem n elementos. Então $m = n$.

Dizemos que um espaço vetorial V tem *dimensão finita n* ou que é *n-dimensional*, e escrevemos

$$\dim V = n$$

se V possui alguma base com n elementos. O Teorema 4.12 nos diz que todas as bases de V possuem o mesmo número de elementos, portanto, esse conceito está bem definido.

O espaço vetorial $\{0\}$ tem dimensão 0, por definição.

Suponha que um espaço vetorial V não possua uma base finita. Então dizemos que V tem *dimensão infinita*.

O Teorema 4.12 fundamental é uma consequência do "lema de substituição" seguinte (demonstrado no Problema 4.35).

Lema 4.13 Se $\{v_1, v_2, ..., v_n\}$ gera V e $\{w_1, w_2, ..., w_m\}$ é linearmente independente, então $m \leq n$, e V é gerado por um conjunto da forma

$$\{w_1, w_2, \ldots, w_m, \ v_{i_1}, v_{i_2}, \ldots, v_{i_{n-m}}\}$$

Assim, em particular, $n + 1$ ou mais vetores de V são linearmente dependentes.

Observe, no lema acima, que substituindo m dos vetores do conjunto gerador de V por m vetores linearmente independentes continuamos com um conjunto gerador.

Exemplo de bases

Nesta subseção apresentamos exemplos importantes de bases de alguns dos principais espaços vetoriais deste texto.

(a) Espaço vetorial K^n Considere os n vetores seguintes de K^n.

$$e_1 = (1, 0, 0, 0, \dots, 0, 0), \; e_2 = (0, 1, 0, 0, \dots, 0, 0), \; \dots, \; e_n = (0, 0, 0, 0, \dots, 0, 1)$$

Esses vetores são linearmente independentes. (Por exemplo, constituem uma matriz escalonada.) Além disso, qualquer vetor $u = (a_1, a_2, \dots, a_n)$ de K^n pode ser escrito como uma combinação linear desses vetores. Mais precisamente,

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n$$

Por isso, esses vetores formam uma base de K^n, denominada *base canônica* de K^n. Assim (como era de se esperar), K^n tem dimensão n. Em particular, qualquer outra base de K^n tem n elementos.

(b) Espaço vetorial $\mathbf{M} = \mathbf{M}_{r,s}$ de todas as matrizes $r \times s$ As seis matrizes seguintes constituem uma base do espaço vetorial $\mathbf{M}_{2,3}$ de todas as matrizes 2×3 sobre K.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Mais geralmente, no espaço vetorial $\mathbf{M} = \mathbf{M}_{r,s}$ de todas as matrizes $r \times s$, seja E_{ij} a matriz com ij-ésima entrada 1 e zeros nas demais entradas. Então todas essas matrizes formam uma base de $\mathbf{M}_{r,s}$, denominada *base canônica* de $\mathbf{M}_{r,s}$. Assim, $\mathbf{M}_{r,s} = rs$.

(c) Espaço vetorial $\mathbf{P}_n(t)$ de todos os polinômios de grau $\leq n$ O conjunto $S = \{1, t, t^2, t^3, \dots, t^n\}$ de $n + 1$ polinômios é uma base de $\mathbf{P}_n(t)$. Mais precisamente, qualquer polinômio $f(t)$ de grau $\leq n$ pode ser escrito como uma combinação linear dessas potências de t, e podemos mostrar que esses polinômios são linearmente independentes. Assim, dim $\mathbf{P}_n(t) = n + 1$.

(d) Espaço vetorial $\mathbf{P}(t)$ de todos os polinômios Considere qualquer conjunto finito $S = \{f_1(t), f_2(t), \dots, f_m(t)\}$ de polinômios de $\mathbf{P}(t)$ e seja m o maior dos graus desses polinômios. Então, qualquer polinômio $g(t)$ de grau maior do que m não pode ser escrito como uma combinação linear dos elementos de S. Assim, S não pode ser uma base de $\mathbf{P}(t)$. Isso significa que a dimensão de $\mathbf{P}(t)$ é infinita. Observamos que o conjunto infinito $S' = \{1, t, t^2, t^3, \dots\}$, consistindo em todas as potências de t, gera $\mathbf{P}(t)$ e é linearmente independente. Por isso, S' é uma base infinita de $\mathbf{P}(t)$.

Teoremas sobre bases

Os três teoremas seguintes (demonstrados nos Problemas 4.37, 4.38 e 4.39) são muito utilizados.

Teorema 4.14 Se V for um espaço vetorial de dimensão finita n, valem as afirmações seguintes.

(i) Quaisquer $n + 1$ ou mais vetores de V são linearmente dependentes.
(ii) Qualquer conjunto linearmente independente $S = \{u_1, u_2, ..., u_n\}$ com n elementos de V é uma base de V.
(iii) Qualquer conjunto gerador $T = \{v_1, v_2, ..., v_n\}$ de V com n elementos é uma base de V.

Teorema 4.15 Se S for um conjunto gerador de um espaço vetorial V, valem as afirmações seguintes.

(i) Qualquer subconjunto linearmente independente máximo de S constitui uma base de V.
(ii) Excluindo de S cada vetor que for uma combinação linear dos vetores de S que o precedem, os vetores que sobram em S constituem uma base de V.

Teorema 4.16 Sejam V um espaço vetorial de dimensão finita e $S = \{u_1, u_2, ..., u_r\}$ um conjunto de vetores linearmente independentes de V. Então S é parte de uma base de V, ou seja, S pode ser estendido até uma base de V.

Exemplo 4.11

(a) Os quatro vetores de $\mathbf{R}^4$ seguintes constituem uma matriz escalonada.

$$(1, 1, 1, 1), \ (0, 1, 1, 1), \ (0, 0, 1, 1), \ (0, 0, 0, 1)$$

Assim, os vetores são linearmente independentes e, como $\mathbf{R}^4 = 4$, os quatro vetores formam uma base de $\mathbf{R}^4$.

(b) Os $n + 1$ polinômios de $\mathbf{P}_n(t)$ seguintes são de graus crescentes.

$$1, \ t - 1, \ (t - 1)^2, \ \ldots, \ (t - 1)^n$$

Portanto, nenhum dos polinômios é uma combinação linear de polinômios que o precedem. Logo, esses polinômios são linearmente independentes. Além disso, formam uma base de $\mathbf{P}_n(t)$, pois $\dim \mathbf{P}_n(t) = n + 1$.

(c) Considere quatro vetores quaisquer de $\mathbf{R}^3$, digamos,

$$(257, -132, 58), \qquad (43, 0, -17), \qquad (521, -317, 94), \qquad (328, -512, -731)$$

Pelo Teorema 4.14(i), os quatro vetores necessariamente são linearmente dependentes, por serem elementos do espaço vetorial tridimensional $\mathbf{R}^3$.

Dimensão e subespaços

O teorema seguinte (demonstrado no Problema 4.40) fornece a relação básica entre a dimensão de um espaço vetorial e a dimensão de um subespaço.

Teorema 4.17 Seja W um subespaço do espaço vetorial V de dimensão n. Então $W \leq n$. Em particular, se $\dim W = n$, então $W = V$.

Exemplo 4.12 Seja W um subespaço do espaço real $\mathbf{R}^3$. Observe que $\mathbf{R}^3 = 3$. O Teorema 4.17 nos diz que a dimensão de W só pode ser 0, 1, 2 ou 3. Vale o seguinte.

(a) Se $\dim W = 0$, então $W = \{0\}$, um ponto.
(b) Se $\dim W = 1$, então W é uma reta pela origem O.
(c) Se $\dim W = 2$, então W é um plano pela origem O.
(d) Se $\dim W = 3$, então W é o espaço $\mathbf{R}^3$ todo.

4.9 APLICAÇÕES A MATRIZES, POSTO DE UMA MATRIZ

Seja A uma matriz $m \times n$ qualquer sobre um corpo K. Lembre que as linhas de A podem ser vistas como vetores de K^n e que o espaço linha de A, denotado por lin(A), é o subespaço de K^n gerado pelas linhas de A. Temos a seguinte definição.

DEFINIÇÃO O *posto* de uma matriz A, denotado por pos(A), é igual ao número máximo de linhas linearmente independentes de A ou, equivalentemente, igual à dimensão do espaço linha de A.

Por outro lado, lembre que as colunas de uma matriz $m \times n$ qualquer A podem ser vistas como vetores de K^m e que o espaço coluna de A, denotado por col(A), é o subespaço de K^m gerado pelas colunas de A. Embora m possa não ser igual a n, ou seja, as linhas e as colunas de A podem pertencer a espaços vetoriais diferentes, temos o resultado fundamental seguinte.

Teorema 4.18 O número máximo de linhas linearmente independentes de uma matriz A qualquer é igual ao número máximo de colunas linearmente independentes de A. Assim, a dimensão do espaço linha de A é igual à dimensão do espaço coluna de A.

Desta forma, podemos reformular a definição dada de posto de A usando colunas em vez de linha.

Problemas de encontrar bases

Nesta subseção mostramos como uma forma escalonada de uma matriz A qualquer pode nos dar a solução para certos problemas sobre a própria matriz A. Mais precisamente, sejam A e B as matrizes seguintes, em que a matriz escalonada B (cujos pivôs aparecem circulados) é uma forma escalonada de A.

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 5 & 5 & 6 & 4 & 5 \\ 3 & 7 & 6 & 11 & 6 & 9 \\ 1 & 5 & 10 & 8 & 9 & 9 \\ 2 & 6 & 8 & 11 & 9 & 12 \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} \textcircled{1} & 2 & 1 & 3 & 1 & 2 \\ 0 & \textcircled{1} & 3 & 1 & 2 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Os quatro problemas a resolver, relativos à matriz A de colunas C_1, C_2, ..., C_6, são os seguintes.

(a) Encontrar uma base do espaço linha de A.
(b) Encontrar cada coluna de A que seja uma combinação linear das colunas de A que a antecedem.
(c) Encontrar uma base do espaço coluna de A.
(d) Encontrar o posto de A.

(a) É dado que A e B são equivalentes por linhas, portanto, têm o mesmo espaço linha. Além disso, B está em forma escalonada, portanto, suas linhas não nulas são linearmente independentes e, por isso, formam uma base do espaço linha de B. Assim, elas também formam uma base do espaço linha de A. Segue que

$$\text{base de lin}(A): \quad (1, 2, 1, 3, 1, 2), \qquad (0, 1, 3, 1, 2, 1), \qquad (0, 0, 0, 1, 1, 2)$$

(b) Seja $M_k = [C_1, C_2, ..., C_k]$ a submatriz de A que consiste nas primeiras k colunas de A. Então M_{k-1} e M_k são, respectivamente, a matriz de coeficientes e a matriz aumentada da equação vetorial

$$x_1C_1 + x_2C_2 + \cdots + x_{k-1}C_{k-1} = C_k$$

O Teorema 3.9 nos diz que o sistema tem uma solução ou, equivalentemente, que C_k é uma combinação linear das colunas precedentes de A se, e só se, $\text{pos}(M_k) = \text{pos}(M_{k-1})$, onde $\text{pos}(M_k)$ significa o número de pivôs na forma escalonada de M_k. Agora, as primeiras k colunas da matriz escalonada B também constituem uma forma escalonada de M_k. Por isso,

$$\text{pos}(M_2) = \text{pos}(M_3) = 2 \qquad \text{e} \qquad \text{pos}(M_4) = \text{pos}(M_5) = \text{pos}(M_6) = 3$$

Assim, as linhas C_3, C_5 e C_6 são, cada uma, combinação linear das colunas de A que as antecedem.

(c) O fato de que as colunas C_1, C_2 e C_4 não são combinações lineares das colunas que as precedem também nos mostra que são linearmente independentes. Assim, elas formam uma base do espaço coluna de A. Segue que

$$\text{base de col}(A): \quad [1, 2, 3, 1, 2]^T, \qquad [2, 5, 7, 5, 6]^T, \qquad [3, 6, 11, 8, 11]^T$$

Observe que C_1, C_2 e C_4 também podem ser caracterizadas como aquelas colunas de A que contêm os pivôs de qualquer forma escalonada de A.

(d) Observamos, aqui, que três possíveis definições do posto de A fornecem o mesmo valor.

(i) Existem três pivôs em B, que é uma forma escalonada de A.
(ii) Os três pivôs de B correspondem às linhas não nulas de B, que formam uma base do espaço linha de A.
(iii) Os três pivôs de B correspondem às colunas de A, que formam uma base do espaço coluna de A.

Assim, $\text{pos}(A) = 3$.

Aplicação para encontrar uma base de $W = \text{ger}(u_1, u_2, ..., u_r)$

Frequentemente, nos é dada uma lista $S = \{u_1, u_2, ..., u_r\}$ de vetores de K^n e queremos encontrar uma base do subespaço W de K^n gerado pelos vetores dados, ou seja, uma base de

$$W = \text{ger}(S) = \text{ger}(u_1, u_2, ..., u_r)$$

Os dois algoritmos seguintes, essencialmente descritos na subseção precedente, encontram uma tal base (e, portanto, a dimensão) de W.

Algoritmo 4.1 (Algoritmo do espaço linha)

Passo 1 Construa a matriz M cujas *linhas* sejam os vetores dados.

Passo 2 Reduza M à forma escalonada.

Passo 3 A resposta são as linhas não nulas da matriz escalonada.

Às vezes queremos encontrar uma base que utiliza apenas os vetores dados originalmente. O próximo algoritmo realiza esta tarefa.

Algoritmo 4.2 (Algoritmo de eliminação)

Passo 1 Construa a matriz M cujas *colunas* sejam os vetores dados.

Passo 2 Reduza M à forma escalonada.

Passo 3 Para cada coluna C_k da matriz escalonada que não tenha pivô, elimine o vetor u_k da lista S dada.

Passo 4 A resposta são os vetores não eliminados de S (que correspondem às colunas com pivôs).

Enfatizamos que, no primeiro algoritmo, formamos uma matriz cujas linhas são os vetores dados, enquanto no segundo algoritmo formamos uma matriz cujas colunas são os vetores dados.

Exemplo 4.13 Seja W o subespaço de $\mathbf{R}^5$ gerado pelos vetores seguintes.

$$u_1 = (1, 2, 1, 3, 2), \qquad u_2 = (1, 3, 3, 5, 3), \qquad u_3 = (3, 8, 7, 13, 8)$$
$$u_4 = (1, 4, 6, 9, 7), \qquad u_5 = (5, 13, 13, 25, 19)$$

Encontre uma base de W dentre os vetores dados e encontre dim W.

Construímos a matriz M cujas colunas são os vetores dados e encontramos sua forma escalonada.

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 2 & 3 & 8 & 4 & 13 \\ 1 & 3 & 7 & 6 & 13 \\ 3 & 5 & 13 & 9 & 25 \\ 2 & 3 & 8 & 7 & 19 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Os pivôs da matriz escalonada aparecem nas colunas C_1, C_2, C_4. Por isso, eliminamos os vetores u_3 e u_5 dos cinco vetores originais. Os vetores restantes, u_1, u_2, u_4, que correspondem às colunas com pivôs da matriz escalonada, formam uma base de W. Assim, em particular, dim $W = 3$.

OBSERVAÇÃO A justificativa do algoritmo de eliminação dado já foi essencialmente descrita, mas a repetimos aqui para enfatizá-la. A coluna C_3 da matriz escalonada do Exemplo 4.13 não ter um pivô significa que a equação vetorial

$$xu_1 + yu_2 = u_3$$

tem uma solução e, portanto, u_3 é uma combinação linear de u_1 e u_2. Analogamente, a coluna C_5 não ter um pivô significa que u_5 é uma combinação linear dos vetores que o antecedem. Eliminamos cada vetor do conjunto gerador original que seja uma combinação linear dos vetores que o antecedem. Assim, os vetores que não são eliminados formam uma base de W.

Aplicação a sistemas de equações lineares homogêneos

Considere, novamente, um sistema $AX = 0$ de equações lineares homogêneo sobre K com n incógnitas. Pelo Teorema 4.4, o conjunto solução W de um tal sistema é um subespaço de K^n e, portanto, W tem alguma dimensão. Temos o teorema seguinte, cuja demonstração será adiada até o Capítulo 5.

Teorema 4.19 A dimensão do espaço solução W de um sistema homogêneo $AX = 0$ é $n - r$, onde n é o número de incógnitas e r é o posto da matriz de coeficientes A.

No caso em que o sistema $AX = 0$ está em forma escalonada, ele tem exatamente $n - r$ variáveis livres, digamos, $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}}$. Seja v_j a solução obtida tomando $x_{i_j} = 1$ (ou qualquer outra constante não nula) e as demais variáveis livres iguais a 0. Mostramos (Problema 4.50) que as soluções $v_1, v_2, \ldots, v_{n-r}$ são linearmente independentes, portanto, formam uma base do espaço solução W.

Na Seção 3.11, já utilizamos esse processo para encontrar uma base do espaço solução W de um sistema homogêneo $AX = 0$. O Problema 4.50 dá outros três exemplos.

4.10 SOMAS E SOMAS DIRETAS

Sejam U e W dois subconjuntos de um espaço vetorial V. A soma de U e W, denotada por $U + W$, consiste em todas as somas $u + w$, com $u \in U$ e $w \in W$. Isto é,

$$U + W = \{v : v = u + w, \text{ com } u \in U \text{ e } w \in W\}$$

Suponha, agora, que U e W sejam subespaços de V. Então podemos mostrar, facilmente (Problema 4.53) que $U + W$ é um subespaço de V. Lembre que $U \cap W$ também é um subespaço de V. O teorema seguinte (demonstrado no Problema 4.58) relaciona as dimensões desses subespaços.

Teorema 4.20 Sejam U e W subespaços de dimensão finita de um espaço vetorial V. Então $U + W$ tem dimensão finita e

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

Exemplo 4.14 Seja $V = \mathbf{M}_{2,2}$ o espaço vetorial das matrizes 2×2. Seja U o conjunto das matrizes cuja segunda linha é nula e seja W o conjunto das matrizes cuja segunda coluna é nula. Então

$$U = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \right\}, \quad W = \left\{ \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \right\} \quad \text{e} \quad U + W = \left\{ \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} \right\}, \quad U \cap W = \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

Logo, $U + W$ consiste naquelas matrizes cuja entrada inferior direita é zero e $U \cap W$ consiste naquelas matrizes cuja segunda linha e coluna são nulas. Observe que $\dim U = 2$, $\dim W = 2$, $\dim (U \cap W) = 1$. Também $\dim(U + W) = 3$, o que era de se esperar pelo Teorema 4.20. Portanto,

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 2 + 2 - 1 = 3$$

Somas diretas

Dizemos que um espaço vetorial V é a *soma direta* de seus subespaços U e W, e escrevemos

$$V = U \oplus W$$

se cada $v \in V$ puder ser escrito, de modo único, como $v = u + w$, com $u \in U$ e $w \in W$.

O teorema seguinte (demonstrado no Problema 4.59), caracteriza uma tal decomposição.

Teorema 4.21 O espaço vetorial V é a soma direta de seus subespaços U e W se, e só se, (i) $V = U + W$, (ii) $U \cap W = \{0\}$.

Exemplo 4.15 Considere o espaço vetorial $V = \mathbf{R}^3$.

(a) Sejam U o plano xy e W o plano yz, ou seja,

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \quad \text{e} \quad W = \{(0, b, c) : b, c \in \mathbf{R}\}$$

Então $\mathbf{R}^3 = U + W$, pois cada vetor de $\mathbf{R}^3$ é a soma de um vetor de U e um vetor de W. Contudo, $\mathbf{R}^3$ não é a soma direta de U e W, porque tais somas não são únicas. Por exemplo,

$$(3, 5, 7) = (3, 1, 0) + (0, 4, 7) \quad \text{e, também,} \quad (3, 5, 7) = (3, -4, 0) + (0, 9, 7)$$

(b) Sejam U o plano xy e W o eixo z, ou seja,

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \quad \text{e} \quad W = \{(0, 0, c) : c \in \mathbf{R}\}$$

Agora, cada vetor $(a, b, c) \in \mathbf{R}^3$ pode ser escrito como a soma de um vetor de U e um vetor de W de uma única maneira, como segue.

$$(a, b, c) = (a, b, 0) + (0, 0, c)$$

Por isso, $\mathbf{R}^3$ é a soma direta de U e W, ou seja, $\mathbf{R}^3 = U \oplus W$.

Somas diretas gerais

A noção de soma direta pode ser estendida a mais fatores da maneira óbvia. Isto é, V é a *soma direta* dos subespaços $W_1, W_2, ..., W_r$, e escrevemos

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

se cada vetor $v \in V$ puder ser escrito, de modo único, como

$$v = w_1 + w_2 + \cdots + w_r$$

com $w_1 \in W_1, w_2 \in W_2, \ldots, w_r \in W_r$.

Temos os teoremas seguintes.

Teorema 4.22 Suponha que $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ e que, para cada k, S_k seja um subconjunto linearmente independente de W_k. Então valem as afirmações seguintes.

(a) A união $S = \bigcup_k S_k$ é linearmente independente em V.
(b) Se cada S_k é uma base de W_k, então $\bigcup_k S_k$ é uma base de V.
(c) $\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_r$.

Teorema 4.23 Suponha que $V = W_1 + W_2 + \cdots + W_r$ e que $\dim V = \sum_k \dim W_k$. Então $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

4.11 COORDENADAS

Seja V um espaço vetorial n-dimensional sobre K com base $S = \{u_1, u_2, ..., u_n\}$. Então qualquer vetor $v \in V$ pode ser escrito, de modo único, como uma combinação linear dos vetores da base S, digamos,

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

Dizemos que os n escalares $a_1, a_2, ..., a_n$ são as *coordenadas* de v em relação à base S; essas coordenadas formam um vetor $[a_1, a_2, ..., a_n]$ de K^n denominado *vetor de coordenadas* de v em relação à base S. Esse vetor é denotado por $[v]_S$ ou, simplesmente, $[v]$, quando S estiver subentendido. Assim,

$$[v]_S = [a_1, a_2, ..., a_n]$$

Por uma questão de notação, usamos colchetes [. . .] em vez de parênteses (. . .) para denotar o vetor de coordenadas.

OBSERVAÇÃO Os n escalares $a_1, a_2, \ldots, a_n$ também formam o *vetor coluna de coordenadas* $[a_1, a_2, \ldots, a_n]^T$ de v em relação a S. A escolha do vetor coluna em vez do vetor linha para representar as coordenadas de v depende do contexto em que for usado. O uso desses vetores coluna ficará claro no Capítulo 6.

Exemplo 4.16 Considere o espaço vetorial $\mathbf{P}_2(t)$ dos polinômios de grau ≤ 2. Os polinômios

$$p_1 = t + 1, \qquad p_2 = t - 1, \qquad p_3 = (t-1)^2 = t^2 - 2t + 1$$

formam uma base S de $\mathbf{P}_2(t)$. O vetor de coordenadas $[v]$ de $v = 2t^2 - 5t + 9$ em relação a S é obtido como segue. Escrevemos $v = xp_1 + yp_2 + zp_3$ com incógnitas escalares e simplificamos.

$$\begin{aligned}2t^2 - 5t + 9 &= x(t+1) + y(t-1) + z(t^2 - 2t + 1)\\ &= xt + x + yt - y + zt^2 - 2zt + z\\ &= zt^2 + (x + y - 2z)t + (x - y + z)\end{aligned}$$

Em seguida, igualamos os coeficientes de mesmas potências de t para obter o sistema

$$z = 2, \qquad x + y - 2z = -5, \qquad x - y + z = 9$$

A solução desse sistema é $x = 3$, $y = -4$, $z = 2$. Assim,

$$v = 3p_1 - 4p_2 + 2p_3 \text{ e, portanto, } [v] = [3, -4, 2].$$

Exemplo 4.17 Considere o espaço real $\mathbf{R}^3$. Os vetores seguintes formam uma base de $\mathbf{R}^3$.

$$u_1 = (1, -1, 0), \qquad u_2 = (1, 1, 0), \qquad u_3 = (0, 1, 1)$$

As coordenadas de $v = (5, 3, 4)$ em relação à base S são obtidas como segue.

Escrevemos $v = xu_1 + yu_2 + zu_3$, ou seja, escrevemos v como uma combinação linear dos vetores da base usando incógnitas escalares x, y, z. Isso fornece

$$\begin{bmatrix}5\\3\\4\end{bmatrix} = x\begin{bmatrix}1\\-1\\0\end{bmatrix} + y\begin{bmatrix}1\\1\\0\end{bmatrix} + z\begin{bmatrix}0\\1\\1\end{bmatrix}$$

O sistema de equações lineares equivalente é o seguinte.

$$x + y = 5, \qquad -x + y + z = 3, \qquad z = 4$$

A solução do sistema é $x = 3$, $y = 2$, $z = 4$. Assim,

$$v = 3u_1 + 2u_2 + 4u_3 \qquad \text{e, portanto,} \qquad [v]_S = [3, 2, 4].$$

OBSERVAÇÃO 1 Existe uma interpretação geométrica para as coordenadas de um vetor v em relação a uma base S para o espaço real $\mathbf{R}^n$, que ilustramos usando a base S de $\mathbf{R}^3$ do Exemplo 4.17. Inicialmente, considere o espaço $\mathbf{R}^3$ com os eixos x, y, z usuais. Então os vetores da base S determinam um novo sistema de coordenadas de $\mathbf{R}^3$ com eixos, digamos, x', y', z', conforme Figura 4-4, que tem as propriedades seguintes.

(1) O eixo x' tem a direção e o sentido de u_1 com unidade medindo $\|u_1\|$.
(2) O eixo y' tem a direção e o sentido de u_2 com unidade medindo $\|u_2\|$.
(3) O eixo z' tem a direção e o sentido de u_3 com unidade medindo $\|u_3\|$.

Então cada vetor $v = (a, b, c)$ ou, equivalentemente, o ponto $P(a, b, c)$ de $\mathbf{R}^3$ tem novas coordenadas em relação aos novos eixos x', y', z'. Essas novas coordenadas são precisamente $[v]_S$, as coordenadas de v em relação à base S. Assim, conforme mostramos no Exemplo 4.17, as coordenadas do ponto $P(5,3,4)$ em relação aos novos eixos formam o vetor $[3, 2, 4]$.

OBSERVAÇÃO 2 Considere a base canônica $E = \{e_1, e_2, \ldots, e_n\}$ de K^n definida por

$$e_1 = (1, 0, 0, \ldots, 0, 0), \qquad e_2 = (0, 1, 0, \ldots, 0, 0), \qquad \ldots, \qquad e_n = (0, 0, 0, \ldots, 0, 1)$$

$v = (5, 3, 4) = [3, 2, 4]$

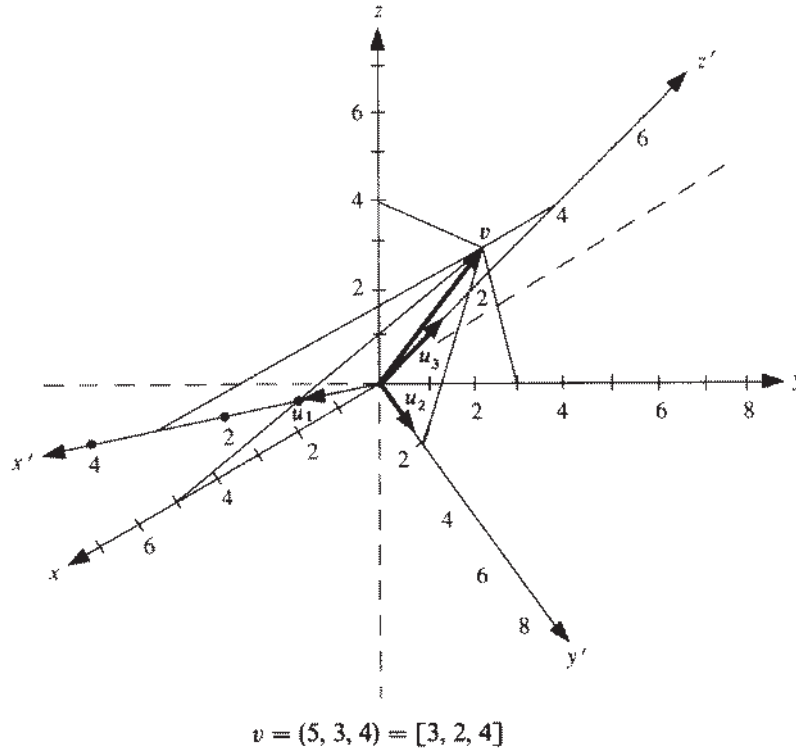



Figura 4-4

Seja $v = (a_1, a_2, ..., a_n)$ um vetor qualquer de K^n. Então é fácil mostrar que

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n, \qquad \text{e, portanto,} \qquad [v]_E = [a_1, a_2, ..., a_n]$$

Assim, o vetor de coordenadas $[v]_E$ de qualquer vetor v em relação à base canônica E de K^n é idêntico ao vetor v original.

Isomorfismo de *V* e K^n

Sejam V um espaço vetorial de dimensão n sobre K e $S = \{u_1, u_2, ..., u_n\}$ uma base de V. Então, a cada vetor $v \in V$ corresponde uma única ênupla $[v]_S$ de K^n. Por outro lado, a cada ênupla $[c_1, c_2, ..., c_n]$ de K^n corresponde um único vetor $c_1u_1 + c_2u_2 + \cdots + c_nu_n$ de V. Assim, a base S induz uma aplicação bijetora entre V e K^n. Além disso, sejam dados

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n \qquad \text{e} \qquad w = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

Então

$$v + w = (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_n + b_n)u_n$$
$$kv = (ka_1)u_1 + (ka_2)u_2 + \cdots + (ka_n)u_n$$

onde k é um escalar. Dessa forma,

$$[v + w]_S = [a_1 + b_1, \quad \ldots, \quad a_n + b_n] = [a_1, \ldots, a_n] + [b_1, \ldots, b_n] = [v]_S + [w]_S$$
$$[kv]_S = [ka_1, ka_2, \ldots, ka_n] = k[a_1, a_2, \ldots, a_n] = k[v]_S$$

Assim, essa aplicação bijetora entre V e K^n preserva as operações de soma de vetores e multiplicação por escalar dos espaços vetoriais. Nesse caso, dizemos que V e K^n são *isomorfos*, e escrevemos

$$V \cong K^n$$

Enunciamos esse resultado formalmente como segue.

Teorema 4.24 Seja V um espaço vetorial n-dimensional sobre um corpo K. Então V e K^n são isomorfos.

O próximo exemplo fornece uma aplicação prática desse teorema.

Exemplo 4.18 Suponha que queiramos determinar se as matrizes seguintes de $V = \mathbf{M}_{2,3}$ são ou não são linearmente dependentes.

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 3 & -4 \\ 6 & 5 & 4 \end{bmatrix}, \qquad C = \begin{bmatrix} 3 & 8 & -11 \\ 16 & 10 & 9 \end{bmatrix}$$

Os vetores de coordenadas das matrizes na base canônica de $\mathbf{M}_{2,3}$ são

$$[A] = [1, 2, -3, 4, 0, 1], \qquad [B] = [1, 3, -4, 6, 5, 4], \qquad [C] = [3, 8, -11, 16, 10, 9]$$

Formamos a matriz M cujas linhas são os vetores de coordenadas obtidos e reduzimos M à forma escalonada.

$$M = \begin{bmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 1 & 3 & -4 & 6 & 5 & 4 \\ 3 & 8 & -11 & 16 & 10 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 2 & -2 & 4 & 10 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Como a matriz escalonada possui apenas duas linhas não nulas, os vetores de coordenadas $[A]$, $[B]$, $[C]$ geram um subespaço bidimensional, portanto, são linearmente dependentes. Por isso, as matrizes A, B, C originais são linearmente dependentes.

Problemas Resolvidos

Espaços vetoriais, combinações lineares

4.1 Sejam u e v dois vetores de um espaço vetorial V. Simplifique cada uma das expressões seguintes.

(a) $E_1 = 3(2u - 4v) + 5u + 7v$, (c) $E_3 = 2uv + 3(2u + 4v)$

(b) $E_2 = 3u - 6(3u - 5v) + 7u$, (d) $E_4 = 5u - \dfrac{3}{v} + 5u$

Multiplicamos e agrupamos os termos.

(a) $E_1 = 6u - 12v + 5u + 7v = 11u - 5v$

(b) $E_2 = 3u - 18u + 30v + 7u = -8u + 30v$

(c) E_3 não está definido porque não definimos o produto de dois vetores.

(d) E_4 não está definido porque não definimos a divisão por um vetor.

4.2 Demonstre o Teorema 4.1. Seja V um espaço vetorial sobre um corpo K.

(i) $r0 = 0$. (ii) $0u = 0$. (iii) Se $ru = 0$, então $r = 0$ ou $u = 0$. (iv) $(-r)u = r(-u) = -ru$.

(i) Pelo Axioma [A_2], com $u = 0$, temos $0 + 0 = 0$. Logo, pelo Axioma [M_1], temos

$$r0 = r(0 + 0) = r0 + r0$$

Somando $-r0$ a ambos lados, obtemos o resultado esperado.

(ii) Para escalares, $0 + 0 = 0$. Logo, pelo Axioma [M_2], temos

$$0u = (0 + 0)u = 0u + 0u$$

Somando $-0u$ a ambos lados, obtemos o resultado esperado.

(iii) Supondo que $ru = 0$ e $r \neq 0$, podemos tomar o escalar r^{-1} tal que $r^{-1}r = 1$. Assim,

$$u = 1u = (r^{-1}r)u = r^{-1}(ru) = r^{-1}0 = 0$$

(iv) Usando $u + (-u) = 0$ e $r + (-r) = 0$, obtemos

$$0 = r0 = r[u + (-u)] = ru + r(-u) \qquad \text{e} \qquad 0 = 0u = [r + (-r)]u = ru + (-r)u$$

Somando $-ru$ a ambos lados da primeira equação, obtemos $-ru = r(-u)$ e, somando $-ru$ a ambos lados da segunda equação, obtemos $-ru = (-r)u$. Assim, $(-r)u = r(-u) = -ru$.

4.3 Mostre que (a) $r(u - v) = ru - rv$, (b) $u + u = 2u$)

(a) Usando a definição de subtração, a saber, que $u - v = u + (-v)$, e o Teorema 4.1(iv), a saber, que $r(-v) = -rv$, obtemos

$$r(u - v) = r[u + (-v)] = ru + r(-v) = ru + (-rv) = ru - rv$$

(b) Usando primeiro o Axioma [M_4] e, depois, o Axioma [M_2], resulta

$$u + u = 1u + 1u = (1 + 1)u = 2u$$

4.4 Escreva $v = (1, -2, 5)$ de $\mathbf{R}^3$ como uma combinação linear dos vetores

$$u_1 = (1, 1, 1),\ u_2 = (1, 2, 3),\ u_3 = (2, -1, 1)$$

Queremos encontrar escalares x, y z, ainda incógnitos, tais que $v = xu_1 + yu_2 + zu_3$. Ou seja, exigimos

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = x\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + z\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned}$$

(Por conveniência, escrevemos os vetores de $\mathbf{R}^3$ como colunas, porque dessa forma é mais fácil encontrar o sistema de equações lineares equivalente.) Reduzindo o sistema à forma escalonada, obtemos o sistema triangular

$$x + y + 2z = 1, \qquad y - 3z = -3, \qquad 5z = 10$$

O sistema é consistente e tem uma solução. Substituindo para trás, obtemos a solução $x = -6$, $y = 3$, $z = 2$. Assim, $v = -6u_1 + 3u_2 + 2u_3$.

Alternativamente, escrevemos a matriz aumentada M do sistema de equações lineares equivalente, em que u_1, u_2, u_3 são as três primeiras colunas de M e v é a última coluna, e reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & -1 & -2 \\ 1 & 3 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -3 & -3 \\ 0 & 2 & -1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & 5 & 10 \end{bmatrix}$$

A última matriz corresponde a um sistema triangular, que tem solução. Substituindo o sistema triangular para trás, obtemos a solução $x = -6$, $y = 3$, $z = 2$. Assim, $v = -6u_1 + 3u_2 + 2u_3$.

4.5 Escreva $v = (2, -5, 3)$ de $\mathbf{R}^3$ como uma combinação linear dos vetores

$$u_1 = (1, -3, 2),\ u_2 = (2, -4, -1),\ u_3 = (1, -5, 7)$$

Queremos encontrar escalares x, y z, ainda incógnitos, tais que $v = xu_1 + yu_2 + zu_3$. Ou seja, exigimos

$$\begin{bmatrix} 2 \\ -5 \\ 3 \end{bmatrix} = x\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ -4 \\ -1 \end{bmatrix} + z\begin{bmatrix} 1 \\ -5 \\ 7 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ -3x - 4y - 5z &= -5 \\ 2x - y + 7z &= 3 \end{aligned}$$

Reduzindo o sistema à forma escalonada, obtemos o sistema

$$x + 2y + z = 2, \qquad 2y - 2z = 1, \qquad 0 = 3$$

O sistema é inconsistente e não tem solução. Assim, v não pode ser escrito como uma combinação linear de u_1, u_2, u_3.

4.6 Escreva o polinômio $v = t^2 + 4t - 3$ de $\mathbf{P}(t)$ como uma combinação linear dos polinômios

$$p_1 = t^2 - 2t + 5, \qquad p_2 = 2t^2 - 3t, \qquad p_3 = t + 1$$

Consideremos v como uma combinação linear de p_1, p_2, p_3 usando incógnitas x, y, z para escrever

$$t^2 + 4t - 3 = x(t^2 - 2t + 5) + y(2t^2 - 3t) + z(t + 1) \qquad (*)$$

Podemos continuar de duas maneiras.

Método 1 Expandimos o lado direito de (*) em termos das potências de t como segue.

$$\begin{aligned} t^2 + 4t - 3 &= xt^2 - 2xt + 5x + 2yt^2 - 3yt + zt + z \\ &= (x + 2y)t^2 + (-2x - 3y + z)t + (5x + 3z) \end{aligned}$$

Igualando os coeficientes de mesma potência de t e reduzindo o sistema à forma escalonada, obtemos

$$\begin{aligned} x + 2y &= 1 \\ -2x - 3y + z &= 4 \\ 5x + 3z &= -3 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y &= 1 \\ y + z &= 6 \\ -10y + 3z &= -8 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y &= 1 \\ y + z &= 6 \\ 13z &= 52 \end{aligned}$$

O sistema é consistente e tem uma solução. Substituindo para trás, obtemos a solução $x = -3$, $y = 2$, $z = 4$. Assim, $v = -3p_1 + 2p_2 + 4p_3$.

Método 2 A equação (*) é uma identidade em t, ou seja, a igualdade vale para cada valor de t. Assim, podemos tomar t igual a quaisquer números para obter equações nas incógnitas, como segue.

(a) Tomando $t = 0$ em (*), obtemos a equação $-3 = 5x + z$
(b) Tomando $t = 1$ em (*), obtemos a equação $2 = 4x - y + 2z$
(c) Tomando $t = -1$ em (*), obtemos a equação $-6 = 8x + 5y$

Resolvemos o sistema das três equações para obter, de novo, a solução $x = -3$, $y = 2$, $z = 4$. Assim, $v = -3p_1 + 2p_2 + 4p_3$.

4.7 Escreva M como uma combinação linear das matrizes A, B, C seguintes.

$$M = \begin{bmatrix} 4 & 7 \\ 7 & 9 \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 4 & 5 \end{bmatrix}$$

Consideremos M como uma combinação linear de A, B, C usando incógnitas escalares x, y, z, ou seja, consideramos $M = xA + yB + zC$. Obtemos

$$\begin{bmatrix} 4 & 7 \\ 7 & 9 \end{bmatrix} = x\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + y\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + z\begin{bmatrix} 1 & 1 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} x + y + z & x + 2y + z \\ x + 3y + 4z & x + 4y + 5z \end{bmatrix}$$

Igualando entradas correspondentes, formamos o sistema de equações equivalente

$$x + y + z = 4, \qquad x + 2y + z = 7, \qquad x + 3y + 4z = 7, \qquad x + 4y + 5z = 9$$

Reduzindo o sistema à forma escalonada, obtemos

$$x + y + z = 4, \qquad y = 3, \qquad 3z = -3, \qquad 4z = -4$$

A última equação é redundante e, portanto, eliminada. Substituindo para trás, obtemos a solução $z = -1$, $y = 3$, $x = 2$. Assim, $M = 2A + 3B - C$.

Subespaços

4.8 Demonstre o Teorema 4.2. W é um subespaço de V se as duas condições seguintes são verdadeiras.

(a) $0 \in W$.

(b) Se $u, v \in W$ e $r \in K$, então $u + v, ru \in W$.

Por (a), W é não vazio e, por (b), as operações de adição vetorial e multiplicação de vetor por escalar estão bem definidas em W. Os Axiomas [A_1], [A_4], [M_1], [M_2], [M_3], [M_4] valem em W porque os vetores de W pertencem a V. Assim, resta mostrar que [A_2] e [A_3] valem em W. Entretanto, [A_2] vale porque o vetor nulo de V pertence a W, por (a). Finalmente, se $v \in W$, então $(-1)v = -v \in W$, e $v + (-v) = 0$. Assim, vale [A_3] em W.

4.9 Seja $V = \mathbf{R}^3$. Mostre que W não é um subespaço, nos casos seguintes.

(a) $W = \{(a, b, c) : a \geq 0\}$, (b) $W = \{(a, b, c) : a^2 + b^2 + c^2 \leq 1\}$.

Em cada caso, mostre que não vale o Teorema 4.2.

(a) W consiste naqueles vetores com primeira entrada não negativa. Assim, $v = (1, 2, 3)$ pertence a W. Seja $r = -3$. Então $rv = (-3, -6, -9)$ não pertence a W, porque -3 é negativo. Assim, W não é um subespaço de V.

(b) W consiste naqueles vetores cujo comprimento não é maior do que 1. Assim, $u = (1, 0, 0)$ e $v = (0, 1, 0)$ pertencem a W, mas $u + v = (1, 1, 0)$ não pertence a W, porque $1^2 + 1^2 + 0^2 = 2 > 1$. Assim, W não é um subespaço de V.

4.10 Seja $V = \mathbf{P}(t)$ o espaço vetorial dos polinômios reais. Decida se W é, ou não é, um subespaço de V, nos casos seguintes.

(a) W consiste em todos os polinômios de coeficientes inteiros.

(b) W consiste em todos os polinômios de grau ≥ 6 e o polinômio zero.

(c) W consiste em todos os polinômios com somente potências pares de t.

(a) Não é, porque a multiplicação por escalar de polinômios de W nem sempre pertence a W. Por exemplo,

$$f(t) = 3 + 6t + 7t^2 \in W \qquad \text{mas} \qquad \tfrac{1}{2}f(t) = \tfrac{3}{2} + 3t + \tfrac{7}{2}t^2 \notin W$$

(b e c) Sim. Em cada caso, W contém o polinômio nulo e as somas e múltiplos escalares de polinômios de W estão em W.

4.11 Seja V o espaço vetorial das funções $f : \mathbf{R} \to \mathbf{R}$. Mostre que W é um subespaço de V, nos casos seguintes.

(a) $W = \{f(x) : f(1) = 0\}$, todas as funções cujo valor em 1 é 0.

(b) $W = \{f(x) : f(3) = f(1)\}$, todas as funções com mesmo valor em 3 e 1.

(c) $W = \{f(t) : f(-x) = -f(x)\}$, todas as funções *ímpares*.

Denotemos por $\hat{0}$ a função nula, isto é, $\hat{0}(x) = 0$, para cada valor de x.

(a) $\hat{0} \in W$, porque $\hat{0}(1) = 0$. Sejam $f, g \in W$. Então $f(1) = 0$ e $g(1) = 0$ e, dados escalares a e b, temos

$$(af + bg)(1) = af(1) + bg(1) = a0 + b0 = 0$$

Assim, $af + bg \in W$ e, portanto, W é um subespaço de V.

(b) $\hat{0} \in W$, porque $\hat{0}(3) = 0 = \hat{0}(1)$. Sejam $f, g \in W$. Então $f(3) = f(1)$ e $g(3) = g(1)$ e, dados escalares a e b, temos

$$(af + bg)(3) = af(3) + bg(3) = af(1) + bg(1) = (af + bg)(1)$$

Assim, $af + bg \in W$ e, portanto, W é um subespaço de V.

(c) $\hat{0} \in W$, porque $\hat{0}(-x) = 0 = -0 = -\hat{0}(x)$. Sejam $f, g \in W$. Então $f(-x) = -f(x)$ e $g(-x) = -g(x)$ e, dados escalares a e b, temos

$$(af + bg)(-x) = af(-x) + bg(-x) = -af(x) - bg(x) = -(af + bg)(x)$$

Assim, $ab + gf \in W$ e, portanto, W é um subespaço de V.

4.12 Demonstre o Teorema 4.3. A interseção de um número qualquer de subespaços de V é um subespaço de V.

Seja $\{W_i : i \in I\}$ uma coleção de subespaços de V e denotemos $W = \cap(W_i : i \in I)$ Como cada W_i é um subespaço de V, temos $0 \in W_i$, para cada $i \in I$. Portanto, $0 \in W$. Se $u, v \in W$, então $u, v \in W_i$, para cada $i \in I$. Como cada W_i é um subespaço, temos $au + bv \in W_i$, para cada $i \in I$. Portanto, $au + bv \in W$. Assim, W é um subespaço de V.

Espaços gerados

4.13 Mostre que os vetores $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 3)$, $u_3 = (1, 5, 8)$ geram $\mathbf{R}^3$.

Devemos mostrar que qualquer vetor $v = (a, b, c)$ de $\mathbf{R}^3$ é uma combinação linear de u_1, u_2, u_3. Escrevemos $v = xu_1 + yu_2 + zu_3$, ou seja,

$$(a, b, c) = x(1, 1, 1) + y(1, 2, 3) + z(1, 5, 8) = (x + y + z, \quad x + 2y + 5z, \quad x + 3y + 8z)$$

Formamos o sistema equivalente e reduzimos à forma escalonada.

$$\begin{matrix} x+y+z=a \\ x+2y+5z=b \\ x+3y+8z=c \end{matrix} \quad \text{ou} \quad \begin{matrix} x+y+z=a \\ y+4z=b-a \\ 2y+7c=c-a \end{matrix} \quad \text{ou} \quad \begin{matrix} x+y+z=a \\ y+4z=b-a \\ -z=c-2b+a \end{matrix}$$

Esse sistema está em forma escalonada e é consistente; de fato,

$$x=-a+5b-3c, \;\; y=3a-7b+4c, \;\; z=a+2b-c$$

é uma solução. Assim, u_1, u_2, u_3 geram $\mathbf{R}^3$.

4.14 Encontre condições sobre a, b, c tais que $v = (a, b, c)$ de $\mathbf{R}^3$ pertença a $W = \text{ger}(u_1, u_2, u_3)$, sendo

$$u_1 = (1, 2, 0), u_2 = (-1, 1, 2), u_3 = (3, 0, -4)$$

Escrevemos v como combinação linear de u_1, u_2, u_3 usando incógnitas escalares x, y, z, ou seja, escrevemos $v = xu_1 + yu_2 + zu_3$. Isso fornece

$$(a,b,c) = x(1,2,0) + y(-1,1,2) + z(3,0,-4) = (x-y+3z, \quad 2x+y, \quad 2y-4z)$$

Formamos o sistema de equações equivalente e reduzimos à forma escalonada, como segue.

$$\begin{matrix} x-y+3z=a \\ 2x+y \quad\;\; =b \\ 2y-4z=c \end{matrix} \quad \text{ou} \quad \begin{matrix} x-y+3z=a \\ 3y-6z=b-2a \\ 2y-4z=c \end{matrix} \quad \text{ou} \quad \begin{matrix} x-y+3z=a \\ 3y-6z=b-2a \\ 0=4a-2b+3c \end{matrix}$$

O vetor $v = (a, b, c)$ pertence a W se, e só se, o sistema é consistente e esse sistema é consistente se, e só se, $4a - 2b + 3c = 0$. Em particular, observe que u_1, u_2, u_3 não geram todo o espaço $\mathbf{R}^3$.

4.15 Mostre que o espaço vetorial $V = \mathbf{P}(t)$ dos polinômios reais não pode ser gerado por um número finito de polinômios.

Qualquer conjunto finito S de polinômios contém um polinômio de grau máximo, digamos, m. Então o espaço gerado por S não pode conter qualquer polinômio de grau maior do que m. Assim, $\text{ger}(S) \neq V$ para qualquer conjunto finito S.

4.16 Demonstre o Teorema 4.5. Seja S um subconjunto de V. (i) Então ger(S) é um subespaço de V que contém S. (ii) Se W é um subespaço de V que contém S, então $\text{ger}(S) \subseteq W$.

(i) Supondo que S seja vazio, temos ger(S) = $\{0\}$, por definição e, portanto, ger(S) = $\{0\}$ é um subespaço de V e $S \subseteq$. Supondo que S seja não vazio, tomamos $v \in S$. Então $v = 1v \in$ e, portanto, $S \subseteq$ ger (S). Também $0 = 0v \in$ ger (S). Agora, dados $u, w \in$ ger (S), digamos,

$$u = a_1u_1 + \cdots + a_ru_r = \sum_i a_iu_i \qquad \text{e} \qquad w = b_1w_1 + \cdots + b_sw_s = \sum_j b_jw_j$$

com $u_i,\ w_j \in S$ e $a_i,\ b_j \in K$, resulta que

$$u + v = \sum_i a_iu_i + \sum_j b_jw_j \qquad \text{e} \qquad au = a\left(\sum_i a_iu_i\right) = \sum_i aa_iu_i$$

pertencem a ger(S), porque cada um é uma combinação linear de vetores de S. Assim, ger(S) é um subespaço de V.

(ii) Dados $u_1, u_2, \ldots, u_r \in S$, cada u_i pertence a W e, portanto, todos múltiplos $a_1u_1, a_2u_2, \ldots, a_ru_r \in W$ e, então, também a soma $a_1u_1 + a_2u_2 + \cdots + a_ru_r \in W$. Assim, W contém todas combinações lineares de elementos de S, ou seja, $\text{ger}(S) \subseteq W$, conforme afirmação.

Dependência linear

4.17 Decida se u e v são linearmente dependentes, ou não, nos casos seguintes.

(a) $u = (1,2),\ v = (3,-5)$, (c) $u = (1,2,-3),\ v = (4,5,-6)$

(b) $u = (1,-3),\ v = (-2,6)$, (d) $u = (2,4,-8),\ v = (3,6,-12)$

Dois vetores u e v são linearmente dependentes se, e só se, um deles é um múltiplo do outro.

(a) Não são. (b) São, pois $v = -2u$. (c) Não são. (d) São, pois $v = \frac{3}{2}u$.

4.18 Decida se u e v são linearmente dependentes ou não nos casos seguintes.

(a) $u = 2t^2 + 4t - 3,\ v = 4t^2 + 8t - 6$, (b) $u = 2t^2 - 3t + 4,\ v = 4t^2 - 3t + 2$,

(c) $u = \begin{bmatrix} 1 & 3 & -4 \\ 5 & 0 & -1 \end{bmatrix}, v = \begin{bmatrix} -4 & -12 & 16 \\ -20 & 0 & 4 \end{bmatrix}$, (d) $u = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}, v = \begin{bmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$

Dois vetores u e v são linearmente dependentes se, e só se, um deles é um múltiplo do outro.

(a) São, pois $v = 2u$. (b) Não são. (c) Sim, pois $v = -4u$. (d) Não são.

4.19 Decida se os vetores $u = (1, 1, 2)$, $v = (2, 3, 1)$, $w = (4, 5, 5)$ de $\mathbf{R}^3$ são linearmente dependentes ou não.

Método 1 Igualamos a zero uma combinação linear de u, v, w usando incógnitas x, y, z para obter o sistema de equações lineares equivalente, que reduzimos à forma escalonada, como segue.

$$x\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + z\begin{bmatrix} 4 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 2y + 4z &= 0 \\ x + 3y + 5z &= 0 \\ 2x + \ \ y + 5z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + 4z &= 0 \\ y + \ \ z &= 0 \end{aligned}$$

O sistema escalonado só tem duas equações com três incógnitas, de modo que tem uma variável livre e uma solução não nula. Assim, u, v, w são linearmente dependentes.

Método 2 Formamos a matriz A de colunas u, v, w e reduzimos à forma escalonada.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 5 \\ 2 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 1 \\ 0 & -3 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

A terceira coluna não tem um pivô, portanto, o terceiro vetor é uma combinação linear dos dois primeiros vetores u e v. Assim, os vetores são linearmente dependentes. (Observe que a matriz A também é a matriz de coeficientes do sistema do Método 1. Em outras palavras, esse método é essencialmente igual ao primeiro.)

Método 3 Formamos a matriz B de linhas u, v, w e reduzimos à forma escalonada.

$$B = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 4 & 5 & 5 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & -3 \\ 0 & 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

Como a matriz escalonada só tem duas linhas não nulas, os três vetores são linearmente dependentes. (Os três vetores dados geram um espaço bidimensional.)

4.20 Decida se os vetores de $\mathbf{R}^3$ dados são linearmente dependentes ou não nos casos seguintes.

(a) $u_1 = (1, 2, 5),\ u_2 = (1, 3, 1),\ u_3 = (2, 5, 7),\ u_4 = (3, 1, 4)$,

(b) $u = (1, 2, 5),\ v = (2, 5, 1),\ w = (1, 5, 2)$,

(c) $u = (1, 2, 3),\ v = (0, 0, 0),\ w = (1, 5, 6)$.

(a) São, pois quaisquer quatro vetores de $\mathbf{R}^3$ são linearmente dependentes.

(b) Usando o Método 2 do problema precedente, formamos a matriz A de colunas u, v, w e reduzimos à forma escalonada.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 5 \\ 5 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & -9 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 24 \end{bmatrix}$$

Cada coluna tem uma entrada pivô, portanto, nenhum vetor é uma combinação linear dos vetores que o antecedem. Assim, os vetores são linearmente independentes.

(c) Como $0 = (0, 0, 0)$ é um dos vetores, os vetores são linearmente dependentes.

4.21 Mostre que as funções $f(t) = \text{sen}\ t$, $g(t) = \cos t$, $h(t) = t$ de $\mathbf{R}$ em $\mathbf{R}$ são linearmente independentes.

Igualamos à função zero $\mathbf{0}$ uma combinação linear de f, g, h usando incógnitas escalares x, y, z, ou seja, tomamos $xf + yg + zh = \mathbf{0}$ e mostramos que $x = 0, y = 0, z = 0$. Enfatizamos que $xf + yg + zh = \mathbf{0}$ significa que, para cada valor de t, vale $xf(t) + yg(t) + zh(t) = 0$.

Assim, na equação $x \operatorname{sen} t + y \cos t + zt = 0$

(i) tomamos $t = 0$ e obtemos $x(0) + y(1) + z(0) = 0$ ou $y = 0$;
(ii) tomamos $t = \pi/2$ e obtemos $x(1) + y(0) + z\pi/2 = 0$ ou $x + \pi z/2 = 0$;
(iii) tomamos $t = \pi$ e obtemos $x(0) + y(-1) + z(\pi) = 0$ ou $-y + \pi z = 0$.

As três equações só têm a solução nula, ou seja, $x = 0, y = 0, z = 0$. Assim, f, g, h são linearmente independentes.

4.22 Suponha que os vetores u, v, w sejam linearmente independentes. Mostre que os vetores $u + v, u - v, u - 2v + w$ também são linearmente independentes.

Considere $x(u + v) + y(u - v) + z(u - 2v + w) = 0$. Então

$$xu + xv + yu - yv + zu - 2zv + zw = 0$$

ou

$$(x + y + z)u + (x - y - 2z)v + zw = 0$$

Como u, v, w são linearmente independentes, os coeficientes da última equação são todos nulos, portanto,

$$x + y + z = 0, \qquad x - y - 2z = 0, \qquad z = 0$$

A única solução desse sistema homogêneo é $x = 0, y = 0, z = 0$. Assim, $u + v, u - v, u - 2v + w$ são linearmente independentes.

4.23 Mostre que os vetores $u = (1 + i, 2i)$ e $w = (1, 1 + i)$ de $\mathbf{C}^2$ são linearmente dependentes sobre o corpo complexo $\mathbf{C}$ mas linearmente independentes sobre o corpo real $\mathbf{R}$.

Lembre que dois vetores são linearmente dependentes (sobre um corpo K) se, e só se, um deles é um múltiplo do outro (com algum elemento de K). Como

$$(1 + i)w = (1 + i)(1,\ 1 + i) = (1 + i,\ 2i) = u$$

temos que u e v são linearmente dependentes sobre $\mathbf{C}$. Por outro lado, u e v são linearmente independentes sobre $\mathbf{R}$, pois nenhum múltiplo de w pode ser igual a u. Mais precisamente, dado r real, o primeiro componente de $rw = (r, r + ri)$ é, necessariamente, real e, portanto, nunca pode ser igual ao primeiro componente $1 + i$ de u, que é complexo.

Base e dimensão

4.24 Decida se os vetores dados formam uma base de $\mathbf{R}^3$, nos casos seguintes.

(a) $(1, 1, 1), (1, 0, 1)$;
(b) $(1, 2, 3), (1, 3, 5), (1, 0, 1), (2, 3, 0)$;
(c) $(1, 1, 1), (1, 2, 3), (2, -1, 1)$;
(d) $(1, 1, 2), (1, 2, 5), (5, 3, 4)$.

(a e b) Não formam, pois uma base de $\mathbf{R}^3$ deve ter exatamente três elementos, já que $\dim \mathbf{R}^3 = 3$.

(c) Os três vetores formam uma base se, e só se, são linearmente independentes. Por isso, formamos a matriz cujas linhas são os vetores dados e reduzimos a matriz à forma escalonada, como segue.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -3 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$$

A matriz escalonada não tem linhas nulas, portanto, os três vetores são linearmente independentes, ou seja, formam uma base.

(d) Formamos a matriz cujas linhas são os vetores dados e reduzimos a matriz à forma escalonada, como segue.

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 5 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & -2 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

A matriz escalonada tem uma linha nula, portanto, os três vetores são linearmente dependentes, ou seja, não formam uma base de $\mathbf{R}^3$.

4.25 Decida se (1, 1, 1, 1), (1, 2, 3, 2), (2, 5, 6, 4), (2, 6, 8, 5) formam uma base de $\mathbf{R}^4$. Se não formarem, encontre a dimensão do espaço gerado.

Formamos a matriz cujas linhas são os vetores dados e reduzimos à forma escalonada.

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 2 & 5 & 6 & 4 \\ 2 & 6 & 8 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 3 & 4 & 2 \\ 0 & 4 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A matriz escalonada tem uma linha nula, portanto, os quatro vetores são linearmente dependentes e não formam uma base de $\mathbf{R}^4$. Como a matriz escalonada tem três linhas não nulas, os quatro vetores geram um subespaço tridimensional.

4.26 Estenda $\{u_1 = (1, 1, 1, 1), u_2 = (2, 2, 3, 4)\}$ a uma base de $\mathbf{R}^4$.

Começamos formando a matriz de linhas u_1 e u_2 e reduzimos à forma escalonada.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Então vemos que $w_1 = (1, 1, 1, 1)$ e $w_2 = (0, 0, 1, 2)$ geram o mesmo espaço gerado pelos vetores u_1 e u_2. Sejam $u_3 = (0, 1, 0, 0)$ e $u_4 = (0, 0, 0, 1)$. Então w_1, u_3, w_2, u_4 formam uma matriz escalonada, portanto, são linearmente independentes e formam uma base de $\mathbf{R}^4$. Assim, u_1, u_2, u_3, u_4 também formam uma base de $\mathbf{R}^4$.

4.27 Considere o corpo complexo **C**, que contém o corpo real **R**, que contém o corpo racional **Q**. (Assim, **C** é um espaço vetorial sobre **R** e **R** é um espaço vetorial sobre **Q**.)

(a) Mostre que $\{1, i\}$ é uma base de **C** sobre **R**; logo, **C** é um espaço vetorial bidimensional sobre **R**.

(b) Mostre que **R** é um espaço de dimensão infinita sobre **Q**.

(a) Dado qualquer $v \in \mathbf{C}$, temos $v = a + bi = a(1) + b(i)$, com $a, b \in \mathbf{R}$. Logo, $\{1, i\}$ gera **C** sobre **R**. Além disso, se $x(1) + y(i) = 0$ ou $x + yi = 0$, com $x, y \in \mathbf{R}$, então $x = 0$ e $y = 0$. Logo, $\{1, i\}$ é linearmente independente sobre **R**. Assim, $\{1, i\}$ é uma base de **C** sobre **R**.

(b) É sabido que π é um número transcendente, ou seja, π não é uma raiz de polinômio algum sobre **Q**. Assim, dado n, os $n + 1$ números reais $1, \pi, \pi^2, \ldots, \pi^n$ são linearmente independentes sobre **Q**, e **R** não pode ter dimensão n sobre **Q**. Por isso, **R** tem dimensão infinita sobre **Q**.

4.28 Seja $S = \{u_1, u_2, \ldots, u_n\}$ um subconjunto de V. Mostre que são equivalentes as definições A e B seguintes de base de V.

(A) S é linearmente independente e gera V.

(B) Cada $v \in V$ é uma combinação linear única de vetores de S.

Digamos que (A) seja válida. Como S gera V, cada vetor v de V é uma combinação linear de vetores de S. Se tivermos duas combinações, digamos,

$$u = a_1u_1 + a_2u_2 + \cdots + a_nu_n \qquad \text{e} \qquad u = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

então obtemos

$$0 = v - v = (a_1 - b_1)u_1 + (a_2 - b_2)u_2 + \cdots + (a_n - b_n)u_n$$

por subtração. Como os u_i são linearmente independentes, todos os coeficientes dessa última combinação são nulos, ou seja,

$$a_1 - b_1 = 0, \qquad a_2 - b_2 = 0, \qquad \ldots, \qquad a_n - b_n = 0$$

Logo, $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$ e as duas combinações são iguais. Isso mostra que vale (B), pois a representação de v como combinação linear de vetores de S é única. Assim, (A) implica (B).

Digamos que (B) seja válida. Então S gera V. Suponha que

$$0 = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

No entanto, temos

$$0 = 0u_1 + 0u_2 + \cdots + 0u_n$$

Por hipótese, a representação de 0 como combinação linear dos u_i é única. Logo, cada $c_i = 0$ e os u_i são linearmente independentes. Assim, (B) implica (A).

Dimensão e subespaços

4.29 Encontre uma base e a dimensão do subespaço W de $\mathbf{R}^3$ nos casos seguintes.

(a) $W = \{(a, b, c) : a + b + c = 0\}$, (b) $W = \{(a, b, c) : (a = b = c)\}$

(a) Observe que $W \neq \mathbf{R}^3$, já que, por exemplo, $(1, 2, 3) \notin W$. Assim, dim $W < 3$. Observe que $u_1 = (1, 0, -1)$ e $u_2 = (0, 1, -1)$ são dois vetores independentes de W. Assim, dim $W = 2$ e, portanto, u_1 e u_2 formam uma base de W.

(b) O vetor $u = (1, 1, 1) \in W$. Qualquer vetor $w \in W$ é da forma $w = (r, r, r)$. Logo, $w = ru$. Assim, u gera W e dim $W = 1$.

4.30 Seja W o subespaço de $\mathbf{R}^4$ gerado pelos vetores

$$u_1 = (1, -2, 5, -3), \qquad u_2 = (2, 3, 1, -4), \qquad u_3 = (3, 8, -3, -5)$$

(a) Encontre uma base e a dimensão de W. (b) Estenda a base de W a uma base de $\mathbf{R}^4$.

(a) Utilizamos o Algoritmo 4.1 do espaço linha. Formamos a matriz cujas linhas são os vetores dados e reduzimos à forma escalonada.

$$A = \begin{bmatrix} 1 & -2 & 5 & -3 \\ 2 & 3 & 1 & -4 \\ 3 & 8 & -3 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 14 & -18 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As linhas não nulas $(1, -2, 5, -3)$ e $(0, 7, -9, 2)$ da matriz escalonada formam uma base do espaço linha de A e, portanto, de W. Em particular, dim $W = 2$.

(b) Queremos encontrar quatro vetores linearmente independentes que incluam os dois vetores de (a). Os quatro vetores $(1, -2, 5, -3)$, $(0, 7, -9, 2)$, $(0, 0, 1, 0)$ e $(0, 0, 0, 1)$ são linearmente independentes (porque formam uma matriz escalonada) e, portanto, formam uma base de $\mathbf{R}^4$ que estende a base de W.

4.31 Seja W o subespaço de $\mathbf{R}^5$ gerado por $u_1 = (1, 2, -1, 3, 4)$, $u_2 = (2, 4, -2, 6, 8)$ $u_3 = (1, 3, 2, 2, 6)$, $u_4 = (1, 4, 5, 1, 8)$, $u_5 = (2, 7, 3, 3, 9)$. Encontre um subconjunto desses vetores que forme uma base de W.

Aqui utilizamos o Algoritmo 4.2 da eliminação. Formamos a matriz M cujas colunas (e não linhas) são os vetores dados e reduzimos à forma escalonada.

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 & 2 \\ 2 & 4 & 3 & 4 & 7 \\ -1 & -2 & 2 & 5 & 3 \\ 3 & 6 & 2 & 1 & 3 \\ 4 & 8 & 6 & 8 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 6 & 5 \\ 0 & 0 & -1 & -2 & -3 \\ 0 & 0 & 2 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

As posições dos pivôs são nas colunas C_1, C_3 e C_5. Logo, os correspondentes vetores u_1, u_3 e u_5 formam uma base de W e dim $W = 3$.

4.32 Seja V o espaço vetorial das matrizes 2×2 sobre K. Seja W o subespaço das matrizes simétricas. Mostre que dim $W = 3$ encontrando uma base de W.

Lembre que uma matriz $A = [a_{ij}]$ é simétrica se $A^T = A$ ou, equivalentemente, se cada $a_{ij} = a_{ji}$. Assim, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ denota uma matriz 2×2 simétrica arbitrária. Tomando (i) $a = 1, b = 0, d = 0$; (ii) $a = 0, b = 1, d = 0$; (iii) $a = 0, b = 0, d = 1$, obtemos as respectivas matrizes

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Afirmamos que $S = \{E_1, E_2, E_3\}$ é uma base de W, ou seja, (a) S gera W e (b) S é linearmente independente.

(a) A matriz dada satisfaz $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} = aE_1 + bE_2 + dE_3$. Assim, S gera W.

(b) Seja $xE_1 + yE_2 + zE_3 = 0$, com x, y, z incógnitas escalares. Ou seja,

$$x\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + y\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} x & y \\ y & z \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Igualando as entradas correspondentes, obtemos $x = 0$, $y = 0$, $z = 0$. Logo S é linearmente independente. Assim, S é uma base de W, conforme afirmamos.

Teoremas sobre dependência linear, bases e dimensão

4.33 Demonstre o Lema 4.10. Suponha que dois ou mais vetores $v_1, v_2, ..., v_m$ sejam linearmente dependentes. Então algum dos vetores é uma combinação linear dos vetores que o antecedem.

Como os v_i são linearmente dependentes, existem escalares $a_1, ..., a_m$ não todos nulos tais que $a_1v_1 + \cdots + a_mv_m = 0$. Seja k o maior inteiro tal que $a_k \neq 0$. Então

$$a_1v_1 + \cdots + a_kv_k + 0v_{k+1} + \cdots + 0v_m = 0 \quad \text{ou} \quad a_1v_1 + \cdots + a_kv_k = 0$$

Suponha que $k = 1$. Então $a_1v_1 = 0$, $a_1 \neq 0$ e, portanto, $v_1 = 0$. Como os v_i são vetores não nulos, estabelecemos que $k > 1$ e obtemos

$$v_k = -a_k^{-1}a_1v_1 - \cdots - a_k^{-1}a_{k-1}v_{k-1}$$

Isso mostra que v_k é uma combinação dos vetores que o antecedem.

4.34 Seja $S = \{v_1, v_2, ..., v_m\}$ um conjunto gerador do espaço vetorial V.

(a) Se $w \in V$, então $\{w, v_1, \ldots, v_m\}$ é linearmente dependente e gera V.

(b) Se v_i é uma combinação linear de $v_1, \ldots, v_{i-1}$, então S sem v_i ainda gera V.

(a) O vetor w é uma combinação linear dos v_i, porque S gera V. Por isso, $\{w, v_1, \ldots, v_m\}$ é linearmente independente. Certamente, w junto com os v_i ainda gera V, porque os v_i geram V, ou seja, $\{w, v_1, \ldots, v_m\}$ gera V.

(b) Digamos que $v_i = k_1v_1 + \cdots + k_{i-1}v_{i-1}$. Seja $u \in V$. Como S gera V, u é uma combinação linear dos v_j, digamos, $u = a_1v_1 + \cdots + a_mv_m$. Substituindo v_i nessa expressão, obtemos

$$\begin{aligned} u &= a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_i(k_1v_1 + \cdots + k_{i-1}v_{i-1}) + a_{i+1}v_{i+1} + \cdots + a_mv_m \\ &= (a_1 + a_ik_1)v_1 + \cdots + (a_{i-1} + a_ik_{i-1})v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_mv_m \end{aligned}$$

Assim, $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ gera V. Em outras palavras, podemos tirar v_i do conjunto gerador e continuar com um conjunto gerador.

4.35 Demonstre o Lema 4.13. Se $\{v_1, v_2, ..., v_n\}$ gera V e $\{w_1, w_2, ..., w_m\}$ é linearmente independente, então $m \leq n$ e V é gerado por um conjunto da forma

$\{w_1, w_2, \ldots, w_m, \ v_{i_1}, v_{i_2}, \ldots, v_{i_{n-m}}\}$

Assim, $n + 1$ ou mais vetores de V são linearmente dependentes.

Basta provar o lema no caso em que nem todos os v_i são nulos. (Verifique isso!). Como $\{v_i\}$ gera V, o Problema 4.34 garante que

$$\{w_1, v_1, \ldots, v_n\} \tag{1}$$

é linearmente dependente e também gera V. Pelo Lema 4.10, algum dos vetores de (1) é uma combinação linear dos vetores que o antecedem. Esse vetor não pode ser w_1, portanto, deve ser um dos v_i, digamos, v_j. Assim, pelo Problema 4.34, podemos tirar v_j do conjunto gerador (1) e obter o conjunto gerador

$$\{w_1, v_1, \ldots, v_{j-1}, \ v_{j+1}, \ldots, v_n\} \tag{2}$$

Agora, repetimos o argumento com o vetor w_2, ou seja, por (2) gerar V, o conjunto

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, \quad v_{j+1}, \ldots, v_n\} \tag{3}$$

é linearmente dependente e também gera V. Novamente, pelo Lema 4.10, algum dos vetores de (3) é uma combinação linear dos vetores que o antecedem. Enfatizamos que esse vetor não pode ser w_1 ou w_2, porque $\{w_1, ..., w_m\}$ é linearmente independente; portanto, deve ser um dos v_i, digamos, v_k. Assim, pelo Problema 4.34, podemos tirar v_k do conjunto gerador (3) e obter o conjunto gerador

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, \quad v_{j+1}, \ldots, v_{k-1}, \quad v_{k+1}, \ldots, v_n\}$$

Repetimos o argumento com o vetor w_3, e assim por diante. Em cada passo, conseguimos acrescentar um dos vetores w_i e tirar um dos v_i do conjunto gerador. Se $m \leq n$, obtemos dessa maneira um conjunto gerado da forma desejada,

$$\{w_1, \ldots, w_m, \quad v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Finalmente, mostramos que não é possível ter $m > n$. De fato, nesse caso, depois de n passos, obteríamos o conjunto gerador $\{w_1, ..., w_n\}$. Isso implicaria que w_{n+1} é uma combinação linear de $w_1, ..., w_n$, contradizendo a hipótese de independência linear do conjunto $\{w_1, ..., w_m\}$.

4.36 Demonstre o Teorema 4.12. Todas bases de um espaço vetorial V têm o mesmo número de elementos.

Sejam $\{u_1, u_2, ..., u_n\}$ e $\{v_1, v_2, ...\}$ duas bases de V. Como $\{u_i\}$ gera V, a base $\{v_1, v_2, ...\}$ não pode ter mais do que n elementos, pois, caso contrário, seria linearmente dependente pelo Problema 4.35 − Lema 4.13. Por outro lado, se a base $\{v_1, v_2, ...\}$ contivesse menos do que n elementos, então $\{u_1, u_2, ..., u_n\}$ seria linearmente dependente pelo Problema 4.35. Assim, a base $\{v_1, v_2, ...\}$ contém exatamente n elementos, provando o teorema.

4.37 Demonstre o Teorema 4.14. Se V for um espaço vetorial de dimensão finita n, valem as afirmações seguintes.

(i) Quaisquer $n + 1$ ou mais vetores de V são linearmente dependentes.

(ii) Qualquer conjunto linearmente independente $S = \{u_1, u_2, ..., u_n\}$ com n elementos de V é uma base de V.

(iii) Qualquer conjunto gerador $T = \{v_1, v_2, ..., v_n\}$ de V com n elementos é uma base de V.

Seja $B = \{w_1, w_2, ..., w_n\}$ uma base de V.

(i) Como B gera V, quaisquer $n + 1$ vetores são linearmente dependentes pelo Lema 4.13.

(ii) Pelo Lema 4.13, podemos acrescentar a S alguns elementos de B até formar um conjunto gerador de V com n elementos. Como S já tem n elementos, S já e um conjunto gerador de V. Assim, S é uma base de V.

(iii) Suponha que T seja linearmente dependente. Então, algum v_i é uma combinação linear dos vetores que o antecedem. Pelo Problema 4.34, V é gerado pelos vetores de T sem v_i e há $n - 1$ deles. Pelo Lema 4.13, o conjunto linearmente independente B não pode ter mais do que $n - 1$ elementos. Isso contradiz o fato de que B tem n elementos. Assim, T é linearmente independente e, portanto T é uma base de V.

4.38 Demonstre o Teorema 4.15. Se S for um conjunto gerador de um espaço vetorial V, valem as afirmações seguintes.

(i) Qualquer subconjunto linearmente independente máximo de S constitui uma base de V.

(ii) Excluindo de S cada vetor que for uma combinação linear dos vetores de S que o antecedem, os vetores que sobram em S constituem uma base de V.

(i) Suponha que $\{v_1, ..., v_m\}$ seja um subconjunto de S com número máximo de elementos linearmente independentes e consideremos $\{v_1, \ldots, v_m, w\}$. Então, $\{v_1, ..., v_m, w\}$ é linearmente dependente. Como nenhum v_k pode ser uma combinação linear de vetores que o antecedem, necessariamente w é uma combinação linear dos v_i. Logo, $w \in$ ger (v_i) e, portanto, $S \subseteq$ ger (v_i). Isso mostra que

$$V = \text{ger}(S) \subseteq \text{ger}(v_i) \subseteq V$$

Assim, $\{v_i\}$ gera V e, por ser linearmente independente, é uma base de V.

4.39 Demonstre o Teorema 4.16. Sejam V um espaço vetorial de dimensão finita e $S = \{u_1, u_2, ..., u_r\}$ um conjunto de vetores linearmente independentes de V. Então S é parte de uma base de V, ou seja, S pode ser estendido até uma base de V.

Suponha que $B = \{w_1, w_2, ..., w_n\}$ seja uma base de V. Então B gera V e, portanto, V é gerado por

$$S \cup B = \{u_1, u_2, \ldots, u_r, \ \ w_1, w_2, \ldots, w_n\}$$

Pelo Teorema 4.15, podemos tirar de $S \cup B$ cada vetor que seja uma combinação linear de vetores que o antecedem até obter uma base B' de V. Como S é linearmente independente, nenhum u_k é uma combinação linear de vetores que o antecedem. Assim, B' contém cada vetor de S e, portanto, S é parte da base B' de V.

4.40 Demonstre o Teorema 4.17. Seja W um subespaço do espaço vetorial V de dimensão n. Então $W \leq n$. Em particular, se dim $W = n$, então $W = V$.

Como V tem dimensão n, quaisquer $n + 1$ vetores são linearmente dependentes. Além disso, como uma base de W consiste em vetores linearmente independentes, não pode conter mais do que n elementos. Por isso, $W \leq n$.
Em particular, se $\{w_1, ..., w_n\}$ for uma base de W, então, por ser um conjunto linearmente independente de n elementos, também é uma base de V. Assim, $W = V$ quando dim $W = n$.

Posto de uma matriz, espaços linha e coluna

4.41 Encontre o posto e uma base do espaço linha de cada uma das matrizes seguintes.

(a) $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix}$.

(a) Reduzimos A à forma escalonada

$$A \sim \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & -6 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As duas linhas não nulas (1, 2, 0, –1) e (0, 2, –3, –1) da forma escalonada de A formam uma base de lin(A). Em particular, pos(A) = 2.

(b) Reduzimos B à forma escalonada

$$B \sim \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & -3 & -6 & -3 & 3 \\ 0 & -1 & -2 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

As duas linhas não nulas (1, 3, 1, –2, –3) e (0, 1, 2, 1, –1) da forma escalonada de B formam uma base de lin(B). Em particular, pos(B) = 2.

4.42 Mostre que $U = W$, sendo U e W os subespaços seguintes de $\mathbf{R}^3$.

U = ger(u_1, u_2, u_3) = ger{(1, 1, –1), (2, 3, –1), (3, 1, –5)}
W = ger(w_1, w_2, w_3) = ger{(1, –1, –3), (3, –2, –8), (2, 1, –3)}

Formamos a matriz A cujas linhas são os u_i e reduzimos A à forma canônica por linhas.

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & -1 \\ 3 & 1 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Em seguida, formamos a matriz B cujas linhas são os w_j e reduzimos B à forma canônica por linhas.

$$B = \begin{bmatrix} 1 & -1 & -3 \\ 3 & -2 & -8 \\ 2 & 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Como A e B têm a mesma forma canônica por linhas, os espaços linha de A e B são iguais, ou seja, $U = W$.

4.43 Seja $A = \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 2 & 4 & 3 & 7 & 7 & 4 \\ 1 & 2 & 2 & 5 & 5 & 6 \\ 3 & 6 & 6 & 15 & 14 & 15 \end{bmatrix}$.

(a) Encontre pos(M_k), com $k = 1, 2, ..., 6$, sendo M_k a submatriz de A consistindo nas k primeiras colunas $C_1, C_2, ..., C_k$ de A.

(b) Quais das colunas C_{k+1} são combinações lineares das colunas $C_1, ..., C_k$ que as antecedem?

(c) Encontre colunas de A que formam uma base do espaço coluna de A.

(d) Escreva a coluna C_4 como uma combinação linear das colunas da parte (c).

(a) Reduzimos A à forma escalonada, como segue.

$$A \sim \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 1 & 3 & 2 & 5 \\ 0 & 0 & 3 & 9 & 5 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe que simultaneamente reduzimos todas as matrizes M_k à forma escalonada. Por exemplo, as quatro primeiras colunas da forma escalonada de A são uma forma escalonada para M_4. Sabemos que pos(M_k) é igual ao número de pivôs ou, equivalentemente, ao número de linhas não nulas numa forma escalonada de M_k. Assim,

$$\operatorname{pos}(M_1) = \operatorname{pos}(M_2) = 1, \qquad \operatorname{pos}(M_3) = \operatorname{pos}(M_4) = 2$$
$$\operatorname{pos}(M_5) = \operatorname{pos}(M_6) = 3$$

(b) A equação vetorial $x_1C_1 + x_2C_2 + \cdots + x_kC_k = C_{k+1}$ fornece o sistema com matriz de coeficientes M_k e matriz aumentada M_{k+1}. Assim, C_{k+1} é uma combinação linear de $C_1, ..., C_k$ se, e só se, pos(M_k) = pos(M_{k+1}) ou, equivalentemente, se C_{k+1} não contém algum pivô. Assim, cada uma das colunas C_2, C_4 e C_6 é uma combinação linear das colunas que a antecedem.

(c) Na forma escalonada de A, os pivôs estão nas primeira, terceira e quinta colunas. Assim, as colunas C_1, C_3, C_5 de A formam uma base do espaço coluna de A. Alternativamente, eliminando as colunas C_2, C_4, C_6 do conjunto gerador de colunas (são combinações lineares de outras colunas) obtemos, de novo, C_1, C_3, C_5.

(d) A matriz escalonada nos diz que C_4 é uma combinação linear das colunas C_1 e C_3. A matriz aumentada M da equação vetorial $C_4 = xC_1 + yC_3$ consiste nas colunas C_1, C_3, C_4 de A que, reduzida à forma escalonada, fornece a matriz (omitindo as linhas nulas) seguinte.

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \end{bmatrix} \quad \text{ou} \quad \begin{matrix} x + y = 2 \\ y = 3 \end{matrix} \quad \text{ou} \quad x = -1, \quad y = 3$$

Assim, $C_4 = -C_1 + 3C_3 = -C_1 + 3C_3 + 0C_5$.

4.44 Suponha que $u = (a_1, a_2, ..., a_n)$ seja uma combinação linear das linhas $R_1, R_2, ..., R_m$ de uma matriz $B = [b_{ij}]$, digamos, $u = k_1R_1 + k_2R_2 + \cdots + k_mR_m$. Mostre que

$$a_i = k_1b_{1i} + k_2b_{2i} + \cdots + k_mb_{mi}, \qquad i = 1, 2, \ldots, n$$

onde $b_{1i}, b_{2i}, \ldots, b_{mi}$ são as entradas da i-ésima linha de B.

É dado que $u = k_1R_1 + k_2R_2 + \cdots + k_mR_m$. Então,

$$\begin{aligned}(a_1, a_2, \ldots, a_n) &= k_1(b_{11}, \ldots, b_{1n}) + \cdots + k_m(b_{m1}, \ldots, b_{mn}) \\ &= (k_1b_{11} + \cdots + k_mb_{m1}, \ldots, k_1b_{1n} + \cdots + k_mb_{mn})\end{aligned}$$

Igualando entradas correspondentes, obtemos o resultado esperado.

4.45 Demonstre o Teorema 4.7. Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ duas matrizes escalonadas equivalentes por linhas com respectivas entradas pivô

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r} \qquad \text{e} \qquad b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$$

(representadas na Figura 4-5). Então A e B têm o mesmo número de linhas não nulas, ou seja, $r = s$, e as entradas pivô estão nas mesmas posições, ou seja, $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$.

$$A = \begin{bmatrix} a_{1j_1} & * & * & * & * & * & * \\ & & a_{2j_2} & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & a_{rj_r} & * & * \end{bmatrix}, \qquad b = \begin{bmatrix} b_{1k_1} & * & * & * & * & * & * \\ & & b_{2k_2} & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & b_{sk_s} & * & * \end{bmatrix}$$

Figura 4-5

Claramente, $A = 0$ se, e só se, $B = 0$ e, portanto, basta demonstrar o teorema com $r \geq 1$ e $s \geq 1$. Primeiro mostramos que $j_1 = k_1$. Suponha que $j_1 < k_1$. Então a coluna j_1 de B é nula. Como a primeira linha R^* de A está no espaço linha de B, temos $R^* = c_1R_1 + c_1R_2 + \cdots + c_mR_m$, onde R_i são as linhas de B. Como a coluna j_1 de B é nula, temos

$$a_{1j_1} = c_1 0 + c_2 0 + \cdots + c_m 0 = 0$$

Mas isso contradiz o fato de que a entrada pivô $a_{1j_1} \neq 0$. Logo, $j_1 \geq k_1$ e, analogamente, $k_1 \geq j_1$. Assim, $j_1 = k_1$.
Agora, seja A' a submatriz de A obtida suprimindo a primeira linha de A e seja B' a submatriz de B obtida suprimindo a primeira linha de B. Vamos provar que A' e B' têm o mesmo espaço linha. Com isso, o teorema resulta demonstrado por indução, pois A' e B' também são matrizes escalonadas.
Seja $R = (a_1, a_2, \ldots, a_n)$ uma linha qualquer de A' e sejam $R_1, \ldots, R_m$ as linhas de B. Como R está no espaço linha de B, existem escalares $d_1, \ldots, d_m$ tais que $R = d_1R_1 + d_2R_2 + \cdots + d_mR_m$. Como A é uma matriz escalonada e R não é a primeira linha de A, a entrada j_1 de R é nula: $a_i = 0$, com $i = j_1 = k_1$. Além disso, como B é uma matriz escalonada, todas as entradas da coluna k_1 de B são nulas, exceto a primeira: $b_{1k_1} \neq 0$, mas $b_{2k_1} = 0, \ldots, b_{mk_1} = 0$. Logo,

$$0 = a_{k_1} = d_1 b_{1k_1} + d_2 0 + \cdots + d_m 0 = d_1 b_{1k_1}$$

Agora, $b_{1k_1} \neq 0$ e, portanto, $d_1 = 0$. Decorre que R é uma combinação linear de $R_2, \ldots, R_m$ e, portanto, está no espaço linha de B'. Como R foi uma linha qualquer de A', resulta que o espaço linha de A' está contido no espaço linha de B'. Analogamente, o espaço linha de B' está contido no espaço linha de A'. Assim, A' e B' têm o mesmo espaço linha e o teorema resulta demonstrado.

4.46 Demonstre o Teorema 4.8. Sejam A e B matrizes em forma canônica por linhas. Então A e B têm o mesmo espaço linha se, e só se, A e B têm as mesmas linhas não nulas.

Obviamente, se A e B têm as mesmas linhas nulas, então têm o mesmo espaço linha. Logo, basta provar a recíproca.
Suponha que A e B tenham o mesmo espaço linha e suponha que $R \neq 0$ seja a i-ésima linha de A. Então existem escalares $c_1, \ldots, c_s$ tais que

$$R = c_1R_1 + c_2R_2 + \cdots + c_sR_s \tag{1}$$

em que R_i são as linhas não nulas de B. O teorema está provado se mostrarmos que $R = R_i$, ou seja, que $c_i = 1$, mas $c_k = 0$, com $k \neq i$.
Seja a_{ij} a entrada pivô de R, ou seja, a primeira entrada não nula de R. Por (1) e o Problema 4.44,

$$a_{ij_i} = c_1 b_{1j_i} + c_2 b_{2j_i} + \cdots + c_s b_{sj_i} \tag{2}$$

Mas, pelo Problema 4.45, b_{ij_i} é um pivô de B e, como B está em forma canônica por linhas, é a única entrada não nula da j-ésima coluna de B. Assim, de (2), decorre que $a_{ij_i} = c_i b_{ij_i}$. Contudo, $a_{ij_i} = 1$ e $b_{ij_i} = 1$, porque A e B estão em forma canônica por linhas, portanto, $c_i = 1$.
Agora, suponha que $k \neq i$ e seja b_{kj_k} a entrada pivô de R_k. Por (1) e o Problema 4.44,

$$a_{ij_k} = c_1 b_{1j_k} + c_2 b_{2j_k} + \cdots + c_s b_{sj_k} \tag{3}$$

Como B está em forma canônica por linhas, b_{kj_k} é a única entrada não nula da j-ésima coluna de B. Assim, de (3), decorre que $a_{ij_k} = c_k b_{kj_k}$. Além disso, pelo Problema 4.45, a_{kj_k} é uma entrada pivô de A e, como A está em forma canônica por linhas, $a_{ij_k} = 0$. Assim, $c_k b_{kj_k} = 0$ e, como $b_{kj_k} = 1$, resulta $c_k = 0$. Por isso, $R = R_i$, demonstrando o teorema.

4.47 Demonstre o Corolário 4.9. Toda matriz A é equivalente por linhas a uma única matriz em forma canônica por linhas.

Suponha que A seja equivalente por linhas às matrizes A_1 e A_2, ambas em forma canônica por linhas. Então $\text{lin}(A) = \text{lin}(A_1)$ e $\text{lin}(A) = \text{lin}(A_2)$. Logo, $\text{lin}(A_1) = \text{lin}(A_2)$. Como A_1 e A_2 estão em forma canônica por linhas, o Teorema 4.8 garante que $A_1 = A_2$, provando o corolário.

4.48 Suponha que os produtos RB e AB estejam definidos, sendo R um vetor linha e A e B matrizes quaisquer. Demonstre as afirmações seguintes.

(a) RB é uma combinação linear das linhas de B.

(b) O espaço linha de AB está contido no espaço linha de B.

(c) O espaço coluna de AB está contido no espaço coluna de A.

(d) Se C for um vetor coluna e AC estiver definido, então AC é uma combinação linear das colunas de A.

(e) $\text{pos}(AB) \leq \text{pos}(B)$ e $\text{pos}(AB) \leq \text{pos}(A)$.

(a) Suponha que $R = (a_1, a_2, ..., a_m)$ e $B = [b_{ij}]$. Sejam $B_1, ..., B_m$ as linhas de B e $B^1, ..., B^n$ suas colunas. Então

$$\begin{aligned} RB &= (RB^1, RB^2, \ldots, RB^n) \\ &= (a_1b_{11} + a_2b_{21} + \cdots + a_mb_{m1}, \quad \ldots, \quad a_1b_{1n} + a_2b_{2n} + \cdots + a_mb_{mn}) \\ &= a_1(b_{11}, b_{12}, \ldots, b_{1n}) + a_2(b_{21}, b_{22}, \ldots, b_{2n}) + \cdots + a_m(b_{m1}, b_{m2}, \ldots, b_{mn}) \\ &= a_1B_1 + a_2B_2 + \cdots + a_mB_m \end{aligned}$$

Assim, RB é uma combinação linear das linhas de B, como queríamos mostrar.

(b) As linhas de AB são R_iB, em que R_i é a i-ésima linha de A. Logo, pela parte (a), cada linha de AB está no espaço linha de B. Assim, $\text{lin}(AB) \subseteq \text{lin}(B)$, como queríamos mostrar.

(c) Usando a parte (b), temos $\text{col}(AB) = \text{lin}(AB)^T = \text{lin}(B^TA^T) \subseteq \text{lin}(A^T) = \text{col}(A)$.

(d) Decorre de (c), trocando B por C.

(e) O espaço linha de AB está contido no espaço linha de B, logo, $\text{pos}(AB) \leq \text{pos}(B)$. Além disso, o espaço coluna de AB está contido no espaço coluna de A, logo, $\text{pos}(AB) \leq \text{pos}(A)$.

4.49 Seja A uma matriz. Mostre que A é invertível se, e só se, $\text{pos}(A) = n$.

Observe que as linhas da matriz identidade I_n de ordem n são linearmente independentes, porque I_n está em forma escalonada; logo, $\text{pos}(I_n) = n$. Sendo A invertível, temos que A é equivalente por linhas a I_n, portanto, $\text{pos}(A) = n$. Se A não for invertível, então A é equivalente por linhas a uma matriz com alguma linha nula, portanto, $\text{pos}(A) < n$. Assim, A é invertível se, e só se, $\text{pos}(A) = n$.

Aplicações a equações lineares

4.50 Encontre a dimensão e uma base do espaço solução W de cada sistema homogêneo dado.

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ x + 2y + 3z + s + t &= 0 \\ 3x + 6y + 8z + s + 5t &= 0 \end{aligned} \qquad \begin{aligned} x + 2y + z - 2t &= 0 \\ 2x + 4y + 4z - 3t &= 0 \\ 3x + 6y + 7z - 4t &= 0 \end{aligned} \qquad \begin{aligned} x + y + 2z &= 0 \\ 2x + 3y + 3z &= 0 \\ x + 3y + 5z &= 0 \end{aligned}$$

(a) (b) (c)

(a) Reduzimos o sistema à forma escalonada, como segue.

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \\ 2z + 4s - 4t &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \end{aligned}$$

O sistema está em forma escalonada e tem duas equações (não nulas) com cinco incógnitas. Logo, o sistema tem $5 - 2 = 3$ variáveis livres, que são y, s, t. Assim, $\dim W = 3$. Para uma base de W,

(1) tomamos $y = 1, s = 0, t = 0$ e obtemos a solução $v_1 = (-2, 1, 0, 0, 0)$;
(2) tomamos $y = 0, s = 1, t = 0$ e obtemos a solução $v_2 = (5, 0, -2, 1, 0)$;
(3) tomamos $y = 0, s = 0, t = 1$ e obtemos a solução $v_3 = (-7, 0, 2, 0, 1)$.

O conjunto $\{v_1, v_2, v_3\}$ é uma base do espaço solução W.

(b) (Aqui utilizamos o formato matricial do nosso sistema homogêneo.) Reduzimos a matriz de coeficientes A à forma escalonada, como segue.

$$A = \begin{bmatrix} 1 & 2 & 1 & -2 \\ 2 & 4 & 4 & -3 \\ 3 & 6 & 7 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & -2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & -2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Isso corresponde ao sistema

$$\begin{aligned} x + 2y + 2z - 2t &= 0 \\ 2z + \quad t &= 0 \end{aligned}$$

As variáveis livres são y e t, sendo dim $W = 2$. Para uma base de W,

(1) tomamos $y = 1$, $t = 0$ e obtemos a solução $u_1 = (-2, 1, 0, 0)$;
(2) tomamos $y = 0$, $t = 2$ e obtemos a solução $u_2 = (6, 0, -1, 2)$.
Então $\{u_1, u_2\}$ é uma base de W.

(c) Reduzimos a matriz de coeficientes A à forma escalonada, como segue.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 1 & 3 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix}$$

Isso corresponde a um sistema triangular sem variáveis livres. Assim, 0 é a única solução, ou seja, $W = \{0\}$ e dim $W = 0$.

4.51 Encontre um sistema homogêneo cujo espaço solução W seja gerado por

$$\{u_1, u_2, u_3\} = \{(1, -2, 0, 3), \quad (1, -1, -1, 4), \quad (1, 0, -2, 5)\}$$

Seja $v = (x, y, z, t)$. Então $v \in W$ se, e só se, v é uma combinação linear dos vetores $\{u_1, u_2, u_3\}$ que geram W. Por isso, formamos a matriz M cujas primeiras colunas sejam u_1, u_2, u_3 e cuja última coluna seja v e, depois reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & 1 & x \\ -2 & -1 & 0 & y \\ 0 & -1 & -2 & z \\ 3 & 4 & 5 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & x \\ 0 & 1 & 2 & 2x+y \\ 0 & -1 & -2 & z \\ 0 & 1 & 2 & -3x+t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & x \\ 0 & 1 & 2 & 2x+y \\ 0 & 0 & 0 & 2x+y+z \\ 0 & 0 & 0 & -5x-y+t \end{bmatrix}$$

Então v é uma combinação linear de u_1, u_2, u_3 se pos(M) = pos(A), em que A é a submatriz sem a coluna v. Assim, tomamos as duas últimas entradas na quarta coluna do lado direito como sendo nulas para obter o sistema homogêneo procurado, ou seja,

$$\begin{aligned} 2x + y + z \qquad &= 0 \\ 5x + y \qquad - t &= 0 \end{aligned}$$

4.52 Sejam $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ as variáveis livres de um sistema homogêneo de equações lineares com n incógnitas. Seja v_j a solução para a qual $x_{i_j} = 1$ e todas as demais variáveis livres iguais a 0. Mostre que as soluções v_1, v_2, ..., v_k são linearmente independentes.

Seja A a matriz cujas linhas são os v_i. Trocamos entre si as colunas 1 e i_1, depois as colunas 2 e i_2, e assim por diante, até as colunas k e i_k, obtendo a matriz $k \times n$

$$B = [I, C] = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & c_{1,k+1} & \ldots & c_{1n} \\ 0 & 1 & 0 & \ldots & 0 & 0 & c_{2,k+1} & \ldots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & c_{k,k+1} & \ldots & c_{kn} \end{bmatrix}$$

Essa matriz B está em forma escalonada, portanto, suas linhas são independentes e, dessa forma, pos$(B) = k$. Como A e B são equivalentes por colunas, têm o mesmo posto e resulta pos$(A) = k$. Como A tem k linhas, decorre que essas linhas (ou seja, os v_i) são linearmente independentes, como queríamos mostrar.

Somas, somas diretas, interseções

4.53 Sejam U e W subespaços de um espaço vetorial V. Mostre que

(a) $U + W$ é um subespaço de V;

(b) U e W estão contidos em $U + W$;

(c) $U + W$ é o menor subespaço de V que contém U e W, ou seja, $U + W = \text{ger}(U, W)$;

(d) $W + W = W$.

(a) Como U e W são subespaços, $0 \in U$ e $0 \in W$. Logo, $0 = 0 + 0$ pertence a $U + W$. Sejam $v, v' \in U + W$. Então $v = u + w$ e $v' = u' + v'$, com $u, u' \in U$ e $w, w' \in W$. Segue que

$$av + bv' = (au + bu') + (aw + bw') \in U + W$$

Assim, $U + W$ é um subespaço de V.

(b) Seja $u \in U$. Como W é um subespaço, $0 \in W$. Logo $u = u + 0$ pertence a $U + W$. Assim, $U \subseteq U + W$. Analogamente, $W \subseteq U + W$.

(c) Como $U + W$ é um subespaço de V que contém U e W, também contém o espaço gerado por U e W, ou seja, ger $(U, W) \subseteq U + W$.

Por outro lado, se $v \in U + W$, então $v = u + w = 1u + 1w$, com $u \in U$ e $w \in W$. Logo, v é uma combinação linear de elementos de $U \cup W$ e, portanto, $v \in \text{ger}(U, W)$. Assim, $U + W \subseteq \text{ger}(U, W)$
As duas inclusão mostram a igualdade dos conjuntos.

(d) Como W é um subespaço de V, sabemos que W é fechado perante a adição vetorial, ou seja, $W + W \subseteq W$. Pela parte (a), $W \subseteq W + W$. Assim, $W + W = W$.

4.54 Considere os subespaços de $\mathbf{R}^5$ dados a seguir.

$$\begin{aligned} U &= \text{ger}(u_1, u_2, u_3) = \text{ger}\{(1,3,-2,2,3), \quad (1,4,-3,4,2), \quad (2,3,-1,-2,9)\} \\ W &= \text{ger}(w_1, w_2, w_3) = \text{ger}\{(1,3,0,2,1), \quad (1,5,-6,6,3), \quad (2,5,3,2,1)\} \end{aligned}$$

Encontre uma base e a dimensão de (a) $U + W$, (b) $U \cap W$.

(a) $U + W$ é o espaço gerado por todos os seis vetores. Portanto, formamos a matriz cujas linhas são os dados seis vetores e reduzimos à forma escalonada, como segue.

$$\begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & -4 & 4 & 0 \\ 0 & -1 & 7 & -2 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

As linhas não nulas seguintes da matriz escalonada formam uma base de $U \cap W$,

$$(1,3,-2,2,2,3), \qquad (0,1,-1,2,-1), \qquad (0,0,1,0,-1)$$

Assim, $\dim(U + W) = 3$.

(b) Seja $v = (x, y, z, s, t)$ um elemento arbitrário de $\mathbf{R}^5$. Em primeiro lugar, escrevemos dois sistemas homogêneos cujos espaços solução sejam U e W, respectivamente, digamos, como no Problema 4.50.
Formamos a matriz M de colunas u_i e v e reduzimos à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 1 & 2 & x \\ 3 & 4 & 3 & y \\ -2 & -3 & -1 & z \\ 2 & 4 & -2 & s \\ 3 & 2 & 9 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & x \\ 0 & 1 & -3 & -3x + y \\ 0 & 0 & 0 & -x + y + z \\ 0 & 0 & 0 & 4x - 2y + s \\ 0 & 0 & 0 & -6x + y + t \end{bmatrix}$$

Igualando a zero as três últimas entradas da última coluna, obtemos o sistema homogêneo seguinte, cujo espaço solução é U.

$$-x + y + z = 0, \qquad 4x - 2y + s = 0, \qquad -6x + y + t = 0$$

Agora, formamos a matriz M' de colunas w_i e v e reduzimos à forma escalonada, como segue.

$$M' = \begin{bmatrix} 1 & 1 & 2 & x \\ 3 & 5 & 5 & y \\ 0 & -6 & 3 & z \\ 2 & 6 & 2 & s \\ 1 & 3 & 1 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & x \\ 0 & 2 & -1 & -3x+y \\ 0 & 0 & 0 & -9x+3y+z \\ 0 & 0 & 0 & 4x-2y+s \\ 0 & 0 & 0 & 2x-y+t \end{bmatrix}$$

Novamente, igualando a zero as três últimas entradas da última coluna, obtemos o sistema homogêneo seguinte, cujo espaço solução é W.

$$-9+3+z=0, \qquad 4x-2y+s=0, \qquad 2x-y+t=0$$

Combinando ambos sistemas obtidos, resulta um sistema homogêneo de espaço solução igual a $U \cap W$ que reduzimos à forma escalonada, obtendo

$$\begin{aligned} -x+y+\ \ z \qquad\qquad &= 0 \\ 2y+4z+\ \ s \qquad &= 0 \\ 8z+5s+2t &= 0 \\ s-2t &= 0 \end{aligned}$$

Existe uma única variável livre, a saber, t, portanto, $\dim(U \cap W) = 1$. Tomando $t = 2$, obtemos a solução $u =$ (1, 4, –3, 4, 2), que constitui a base procurada de $U \cap W$.

4.55 Suponha que U e W sejam subespaços distintos de dimensão 4 de um espaço vetorial V de dimensão 6. Encontre as possíveis dimensões de $U \cap W$.

Como U e W são distintos, $U + W$ contém U e W propriamente e, portanto, $\dim(U + W) > 4$. Mas $\dim(U + W)$ não pode ser maior do que 6, já que $\dim V = 6$. Logo, temos duas possibilidades: (a) $\dim(U + W) = 5$, ou (b) $\dim(U + W) = 6$. Pelo Teorema 4.20,

$$\dim(U \cap W) = \dim U + \dim W - \dim(U+W) = 8 - \dim(U+W)$$

Assim, (a) $\dim(U \cap W) = 3$ ou (b) $\dim(U \cap W) = 2$.

4.56 Sejam U e W os subespaços de $\mathbf{R}^3$ dados.

$$U = \{(a,b,c) : a = b = c\} \qquad \text{e} \qquad W = \{(0,b,c)\}$$

(Observe que W é o plano yz.) Mostre que $\mathbf{R}^3 = U \oplus W$.

Inicialmente, mostramos que $U \cap W = \{0\}$. Suponha que $v = (a,b,c) \in U \cap W$. Então $a = b = c$ e $a = 0$. Logo, $a = 0$, $b = 0$, $c = 0$. Assim, $v = 0 = (0, 0, 0)$.

Em seguida, mostramos que $\mathbf{R}^3 = U + W$. De fato, se $v = (a,b,c) \in \mathbf{R}^3$, então

$$v = (a,a,a) + (0, b-a, c-a), \qquad \text{com} \qquad (a,a,a) \in U \qquad \text{e} \qquad (0,\ b-a,\ c-a) \in W.$$

Ambas condições $U \cap W = \{0\}$ e $U + W = \mathbf{R}^3$ implicam que $\mathbf{R}^3 = U \oplus W$.

4.57 Suponha que U e W sejam subespaços de um espaço vetorial V, que $S = \{u_i\}$ gere U e que $S' = \{w_j\}$ gere W. Mostre que $S \cup S'$ gera $U + W$. (Desse modo, por indução, se S_i gera W_i, com $i = 1, 2, ..., n$, então $S_1 \cup \ldots \cup S_n$ gera $W_1 + \cdots + W_n$.)

Seja $v \in U + W$. Então $v = u + w$, com $u \in U$ e $w \in W$. Como S gera U, u é uma combinação linear de u_i e, como S' gera W, w é uma combinação linear de w_j, digamos,

$$u = a_1u_{i_1} + a_2u_{i_2} + \cdots + a_ru_{i_r} \qquad \text{e} \qquad v = b_1w_{j_1} + b_2w_{j_2} + \cdots + b_sw_{j_s}$$

em que $a_i, b_j \in K$. Então,

$$v = u + w = a_1u_{i_1} + a_2u_{i_2} + \cdots + a_ru_{i_r} + b_1w_{j_1} + b_2w_{j_2} + \cdots + b_sw_{j_s}$$

Assim, $S \cup S' = \{u_i, w_j\}$ gera $U + W$.

4.58 Demonstre o Teorema 4.20. Sejam U e W subespaços de dimensão finita de um espaço vetorial V. Então $U + W$ tem dimensão finita e

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W)$$

Observe que $U \cap W$ é um subespaço tanto de U quanto de W. Digamos que dim $U = m$, dim $W = n$ e $\dim(U \cap W) = r$. Seja $\{v_1, ...v_r\}$ uma base de $U \cap W$. Pelo Teorema 4.16, podemos estender $\{v_i\}$ a uma base de U e a uma base de W. Digamos que

$$\{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}\} \qquad \text{e} \qquad \{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$$

sejam bases de U e W, respectivamente. Seja

$$B = \{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}, w_1, \ldots, w_{n-r}\}$$

Observe que B tem, exatamente, $m + n - r$ elementos. Assim, o teorema estará demonstrado se conseguimos mostrar que B é uma base de $U + W$. Como $\{v_i, u_j\}$ gera U e $\{v_i, w_k\}$ gera W, a união $B = \{v_i, u_j, w_k\}$ gera $U + W$. Assim, basta mostrar que B é linearmente independente.
Suponha que

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0 \tag{1}$$

em que a_i, b_j, c_k são escalares. Seja

$$v = a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} \tag{2}$$

Por (1), também temos

$$v = -c_1w_1 - \cdots - c_{n-r}w_{n-r} \tag{3}$$

Como $\{v_i, u_j\} \subseteq U$, $v \in U$, temos, por (2); e, como $\{w_k\} \subseteq W$, $v \in W$, temos por (3). Desse modo, $v \in U \cap W$. Agora, $\{v_j\}$ é uma base de $U \cap W$, portanto, existem escalares $d_1, ..., d_r$ para os quais $v = d_1v_1 + \cdots + d_rv_r$. Assim, por (3), temos

$$d_1v_1 + \cdots + d_rv_r + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0$$

Mas $\{v_i, w_k\}$ é uma base de W, portanto, é independente. Logo, a equação precedente força $c_1 = 0, ..., c_{n-r} = 0$. Substituindo esses valores em (1), obtemos

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} = 0$$

Mas $\{v_i, u_j\}$ é uma base de U, portanto, é independente. Logo, a equação precedente força $a_1 = 0, ..., a_r = 0, b_1 = 0, ..., b_{m-r} = 0$.
Como (1) implica que todos os a_i, b_j, c_k são nulos, resulta que $B = \{v_i, u_j, w_k\}$ é independente, demonstrando o teorema.

4.59 Demonstre o Teorema 4.21. $V = U \oplus W$ se, e só se, (i) $V = U + W$, (ii) $U \cap W = \{0\}$.

Suponha que $V = U \oplus W$. Então cada $v \in V$ pode ser escrito de maneira única na forma $v = u + w$, em que $u \in U$ e $w \in W$. Assim, em particular, $V = U + W$. Agora, suponha que $v \in U \cap W$. Então

$$(1)\ v = v + 0, \text{ em que } v \in U,\ \ 0 \in W, \quad (2)\ v = 0 + v, \text{ em que } 0 \in U,\ \ v \in W$$

Assim, $v = 0 + 0 = 0$ e $U \cap W = \{0\}$.
Por outro lado, suponha que $V = U + W$ e que $U \cap W = \{0\}$. Seja $v \in V$. Como $V = U + W$, existem $u \in U$ e $w \in W$ tais que $v = u + w$. Queremos mostrar que tal soma é única. Suponha que também tenhamos $v = u' + w'$, com $u' \in U$ e $w' \in W$. Então

$$u + w = u' + w', \qquad \text{e, portanto,} \qquad u - u' = w' - w$$

Ambos $u - u' \in U$ e $w' - w \in W$, portanto, como $U \cap W = \{0\}$, resulta

$$u - u' = 0, \quad w' - w = 0, \quad \text{e, portanto, } u = u', \quad w = w'.$$

Assim, uma tal soma para $v \in V$ é única e $V = U \oplus W$.

4.60 Demonstre o Teorema 4.22 (para dois fatores). Suponha que $V = U \oplus W$ e que $S = \{u_1, \ldots, u_m\}$ e $S' = \{w_1, \ldots, w_n\}$ sejam subconjuntos linearmente independente de U e W, respectivamente. Então valem as afirmações seguintes.

(a) A união $S \cup S'$ é linearmente independente em V.

(b) Se S e S' forem bases de U e W, respectivamente, então $S \cup S'$ é uma base de V.

(c) $\dim V = \dim U + \dim W$.

(a) Suponha que $a_1u_1 + \cdots + a_mu_m + b_1w_1 + \cdots + b_nw_n = 0$, em que a_i, b_j sejam escalares. Então

$$(a_1u_1 + \cdots + a_mu_m) + (b_1w_1 + \cdots + b_nw_n) = 0 = 0 + 0$$

onde $0, a_1u_1 + \cdots + a_mu_m \in U$ e $0, b_1w_1 + \cdots + b_nw_n \in W$. Como uma tal soma para 0 é única, isso nos leva a

$$a_1u_1 + \cdots + a_mu_m = 0 \qquad \text{e} \qquad b_1w_1 + \cdots + b_nw_n = 0$$

Como S é linearmente independente, cada $a_i = 0$ e, como S' é linearmente independente, cada $b_j = 0$. Assim, $S = S \cup S$ é linearmente independente.

(b) Pela parte (a), $S = S \ \cup S$ é linearmente independente e, pelo Problema 4.55, $S = S \ \cup S'$ gera $V = U + W$. Assim, $S = S \ \cup S$ é uma base de V.

(c) Segue imediatamente da parte (b).

Coordenadas

4.61 Em relação à base $S = \{u_1, u_2\} = \{(1, 1), (2, 3)\}$ de $\mathbf{R}^2$, encontre as coordenadas do vetor v nos casos seguintes. (a) $v = (4, -3)$, (b) $v = (a, b)$.

Em cada caso, escrevemos

$$v = xu_1 + yu_2 = x(1,1) + y(2,3) = (x + 2y,\ x + 3y)$$

e resolvemos em x e y.

(a) Temos

$$(4, -3) = (x + 2y,\ x + 3y) \qquad \text{ou} \qquad \begin{aligned} x + 2y &= 4 \\ x + 3y &= -3 \end{aligned}$$

A solução é $x = 18$, $y = -7$. Assim, $[v] = [18, -7]$.

(b) Temos

$$(a, b) = (x + 2y,\ x + 3y) \qquad \text{ou} \qquad \begin{aligned} x + 2y &= a \\ x + 3y &= b \end{aligned}$$

A solução é $x = 3a - 2b$, $y = -a + b$. Assim, $[v] = [3a - 2b, -a + b]$.

4.62 Encontre o vetor de coordenadas de $v = (a, b, c)$ em $\mathbf{R}^3$ em relação à

(a) base canônica $E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$,

(b) base $S = \{u_1, u_2, u_3\} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$.

(a) Em relação à base canônica E, as coordenadas de $[v]_E$ são iguais às de v, ou seja, $[v]_E = [a, b, c]$.

(b) Escrevemos v como uma combinação linear de u_1, u_2, u_3 usando escalares x, y, z, obtendo

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \text{ou} \qquad \begin{aligned} x + y + z &= a \\ x + y &= b \\ x &= c \end{aligned}$$

Resolvendo o sistema, obtemos $x = c$, $y = b - c$, $z = a - b$. Assim, $[v]_S = [c, b - c, a - b]$.

4.63 Considere o espaço vetorial $\mathbf{P}_3(t)$ dos polinômios de grau ≤ 3.

(a) Mostre que $S = \{(t-1)^3,\ (t-1)^2,\ t-1,\ 1\}$ é uma base de $\mathbf{P}_3(t)$.

(b) Encontre o vetor de coordenadas $[v]$ de $v = 3t^3 - 4t^2 + 2t - 5$ em relação a S.

(a) O grau de $(t - 1)^k$ é k. Escrevendo os polinômios de S em ordem invertida, vemos que nenhum dos polinômios de S é uma combinação linear dos polinômios que o precedem. Assim, os polinômios são linearmente independentes e, como $\dim \mathbf{P}_3(t) = 4$, formam uma base de $\mathbf{P}_3(t)$.

(b) Escrevemos v como uma combinação linear dos vetores da base usando incógnitas escalares x, y, z, s, obtendo

$$\begin{aligned} v = 3t^3 + 4t^2 + 2t - 5 &= x(t-1)^3 + y(t-1)^2 + z(t-1) + s(1) \\ &= x(t^3 - 3t^2 + 3t - 1) + y(t^2 - 2t + 1) + z(t - 1) + s(1) \\ &= xt^3 - 3xt^2 + 3xt - x + yt^2 - 2yt + y + zt - z + s \\ &= xt^3 + (-3x + y)t^2 + (3x - 2y + z)t + (-x + y - z + s) \end{aligned}$$

Em seguida, igualamos os coeficientes de mesma potência t para obter

$$x = 3, \qquad -3x + y = 4, \qquad 3x - 2y + z = 2, \qquad -x + y - z + s = -5$$

Resolvendo o sistema, obtemos $x = 3$, $y = 13$, $z = 19$, $s = 4$. Assim, $[v] = [3, 13, 19, 4]$.

4.64 Encontre o vetor de coordenadas de $A = \begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix}$ no espaço vetorial real $\mathbf{M} = \mathbf{M}_{2,2}$ em relação à

(a) base $S = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$,

(b) base canônica $E = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$.

(a) Escrevemos A como uma combinação linear dos vetores da base usando incógnitas escalares x, y, z, t, como segue.

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix} = x\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + y\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + t\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x+z+t & x-y-z \\ x+y & x \end{bmatrix}$$

Igualando entradas correspondentes, obtemos o sistema

$$x + z + t = 2, \qquad x - y - z = 3, \qquad x + y = 4, \qquad x = -7$$

Resolvendo o sistema, obtemos $x = -7$, $y = 11$, $z = -21$, $t = 30$. Assim, $[A]_S = [-7, 11, -21, 30]$. (Observe que o vetor de coordenadas de A é um vetor de $\mathbf{R}^4$, pois $\dim \mathbf{M} = 4$.)

(b) Escrevendo A como uma combinação linear das matrizes da base, obtemos

$$\begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix} = x\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + y\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + z\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + t\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$$

Assim, $x = 2$, $y = 3$, $z = 4$, $t = -7$. Assim, $[A] = [2, 3, 4, -7]$, cujas entradas são os elementos de A escritos linha por linha.

OBSERVAÇÃO Esse resultado é válido em geral, ou seja, se A for qualquer matriz $m \times n$ de $\mathbf{M} = \mathbf{M}_{m,n}$, então as coordenadas de A em relação à base canônica de $\mathbf{M}$ são os elementos de A escritos linha por linha.

4.65 No espaço $\mathbf{M} = \mathbf{M}_{2,3}$, decida se as matrizes seguintes são linearmente independentes ou não.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 4 & 7 \\ 10 & 1 & 13 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 2 & 5 \\ 8 & 2 & 11 \end{bmatrix}$$

Se as matrizes forem linearmente dependentes, encontre a dimensão e uma base do subespaço W de $\mathbf{M}$ gerado pelas matrizes.

Os vetores de coordenadas das matrizes dadas em relação à base canônica de $\mathbf{M}$ são os seguintes.

$$[A] = [1, 2, 3, 4, 0, 5], \qquad [B] = [2, 4, 7, 10, 1, 13], \qquad [C] = [1, 2, 5, 8, 2, 11]$$

Formamos a matriz M cujas linhas são esses vetores de coordenadas e reduzimos M à forma escalonada.

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 & 0 & 5 \\ 2 & 4 & 7 & 10 & 1 & 13 \\ 1 & 2 & 5 & 8 & 2 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 4 & 0 & 5 \\ 0 & 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Como a matriz escalonada tem somente duas linhas não nulas, os vetores de coordenadas $[A]$, $[B]$, $[C]$ geram um espaço bidimensional e, portanto, são linearmente dependentes. Assim, os vetores A, B, C, são linearmente dependentes. Além disso, $\dim W = 2$ e as matrizes

$$w_1 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \end{bmatrix} \qquad \text{e} \qquad w_2 = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

correspondentes às linhas não nulas da matriz escalonada constituem uma base de W.

Problemas variados

4.66 Considere uma sequência finita de vetores $S = \{v_1, v_2, ..., v_n\}$. Seja T a sequência de vetores obtida a partir de S por uma das "operações elementares" seguintes: (i) trocar de lugar dois vetores, (ii) multiplicar um vetor por um escalar não nulo, (iii) somar um múltiplo de um vetor a um outro vetor. Mostre que S e T geram o mesmo espaço W. Também mostre que T é independente se, e só se, S é independente.

Observe que, em cada operação, os vetores de T são combinações lineares de vetores de S. Por outro lado, cada operação tem uma inversa do mesmo tipo (Prove isso!) e, portanto, os vetores de S são combinações lineares de vetores de T. Assim, S e T geram o mesmo espaço W. Também, T é independente se, e só se, dim $W = n$ e isso é válido se, e só se, S também é independente.

4.67 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes $m \times n$ sobre um corpo K equivalentes por linhas e sejam $v_1, ..., v_n$ vetores quaisquer de um espaço vetorial V sobre K. Sejam

$$\begin{array}{ll} u_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n & w_1 = b_{11}v_1 + b_{12}v_2 + \cdots + b_{1n}v_n \\ u_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & w_2 = b_{21}v_1 + b_{22}v_2 + \cdots + b_{2n}v_n \\ \cdots\cdots\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ u_m = a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n & w_m = b_{m1}v_1 + b_{m2}v_2 + \cdots + b_{mn}v_n \end{array}$$

Mostre que $\{u_i\}$ e $\{w_i\}$ geram o mesmo espaço.

Aplicar uma "operação elementar" do Problema 4.66 a $\{u_i\}$ equivale a aplicar uma operação elementar com as linhas da matriz A. Como A e B são equivalentes por linhas, B pode ser obtida a partir de A por uma sequência de operações elementares com as linhas; assim, $\{w_i\}$ pode ser obtido de $\{u_i\}$ pela sequência correspondente de operações. Dessa forma, $\{u_i\}$ e $\{w_i\}$ geram o mesmo espaço.

4.68 Sejam $v_1, ..., v_n$ vetores de um espaço vetorial V sobre K e seja $P = [a_{ij}]$ uma matriz quadrada de ordem n sobre K. Sejam

$$w_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n, \qquad \ldots, \qquad w_n = a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n$$

(a) Suponha que P seja invertível. Mostre que $\{w_i\}$ e $\{v_i\}$ geral o mesmo espaço; assim, $\{w_i\}$ é independente se, e só se, $\{v_i\}$ é independente.

(b) Suponha que P não seja invertível. Mostre que $\{w_i\}$ é dependente.

(c) Suponha que $\{w_i\}$ seja independente. Mostre que P é invertível.

(a) Como P é invertível, é equivalente por linhas à matriz identidade I. Portanto, pelo Problema 4.67, $\{w_i\}$ e $\{v_i\}$ geram o mesmo espaço. Assim, um conjunto é linearmente independente se, e só se, o outro for.

(b) Como P não é invertível, é equivalente por linhas a uma matriz com alguma linha nula. Isso significa que $\{w_i\}$ gera um espaço que tem um conjunto gerador com menos do que n elementos. Assim, $\{w_i\}$ é dependente.

(c) Essa afirmação é a forma contrapositiva da afirmação (b), portanto, equivale a (b).

4.69 Suponha que $A_1, A_2, ...$ sejam conjuntos de vetores linearmente independentes e que $A_1 \subseteq A_2 \subseteq \ldots$. Mostre que a união $A = A_1 \cup A_2 \cup \ldots$ também é linearmente dependente.

Suponha que A seja linearmente dependente. Então existem vetores $v_1, \ldots, v_n \in A$ e escalares $a_1, \ldots, a_n \in K$, não todos nulos, tais que

$$a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0 \qquad (1)$$

Como $A = \cup A_i$ e os $v_i \in A$, existem conjuntos $A_{i_1}, \ldots, A_{i_n}$ tais que

$$v_1 \in A_{i_1}, \qquad v_2 \in A_{i_2}, \qquad \ldots, \qquad v_n \in A_{i_n}$$

Seja k o maior índice dos conjuntos A_{i_j}, ou seja, $k = \max(i_1, \ldots, i_n)$. Como $A_1 \subseteq A_2 \subseteq \ldots$, segue que cada A_{i_j} está contido em A_k. Logo, $v_1, v_2, \ldots, v_n \in A_k$ e, portanto, por (1), A_k é linearmente dependente, contradizendo nossa hipótese. Assim, A é linearmente independente.

4.70 Seja K um subcorpo de um corpo L e seja L um subcorpo de um corpo E. (Assim, $K \subseteq L \subseteq E$ e K é um subcorpo de E.) Suponha que E seja de dimensão n sobre L e que L seja de dimensão m sobre K. Mostre que E é de dimensão mn sobre K.

Suponha que $\{v_1, ..., v_n\}$ seja uma base de E sobre L e que $\{a_1, ..., a_m\}$ seja uma base de L sobre K. Afirmamos que $\{a_i v_j : i = 1, ..., m, j = 1, ..., n\}$ é uma base de E sobre K. Observe que $\{a_i v_j\}$ contém mn elementos.

Seja w um elemento arbitrário de E. Como $\{v_1, ..., v_n\}$ gera E sobre L, w é uma combinação linear dos v_i com coeficientes em L, como segue.

$$w = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n, \qquad b_i \in L \tag{1}$$

Como $\{a_1, ..., a_m\}$ gera L sobre K, cada $b_i \in L$ é uma combinação linear dos a_j com coeficientes em K, como segue.

$$\begin{aligned} b_1 &= k_{11}a_1 + k_{12}a_2 + \cdots + k_{1m}a_m \\ b_2 &= k_{21}a_1 + k_{22}a_2 + \cdots + k_{2m}a_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_n &= k_{n1}a_1 + k_{n2}a_2 + \cdots + k_{mn}a_m \end{aligned}$$

com $k_{ij} \in K$. Substituindo em (1), obtemos

$$\begin{aligned} w &= (k_{11}a_1 + \cdots + k_{1m}a_m)v_1 + (k_{21}a_1 + \cdots + k_{2m}a_m)v_2 + \cdots + (k_{n1}a_1 + \cdots + k_{nm}a_m)v_n \\ &= k_{11}a_1v_1 + \cdots + k_{1m}a_mv_1 + k_{21}a_1v_2 + \cdots + k_{2m}a_mv_2 + \cdots + k_{n1}a_1v_n + \cdots + k_{nm}a_mv_n \\ &= \sum_{i,j} k_{ji}(a_iv_j) \end{aligned}$$

com $k_{ji} \in K$. Assim, w é uma combinação linear dos $a_i v_j$ com coeficientes em K e, portanto, $\{a_i v_j\}$ gera E sobre K.

A afirmação estará demonstrada assim que provarmos que $\{a_i v_j\}$ é linearmente independente sobre K. Suponha que, para escalares $x_{ji} \in K$, tenhamos $\sum_{i,j} x_{ji}(a_iv_j) = 0$, ou seja,

$$(x_{11}a_1v_1 + x_{12}a_2v_1 + \cdots + x_{1m}a_mv_1) + \cdots + (x_{n1}a_1v_n + x_{n2}a_2v_n + \cdots + x_{nm}a_mv_m) = 0$$

ou

$$(x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m)v_1 + \cdots + (x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m)v_n = 0$$

Como $\{v_1, ..., v_n\}$ é linearmente independente sobre L e os coeficientes dos v_i na última expressão pertencem a L, cada coeficiente deve ser nulo, ou seja,

$$x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m = 0, \qquad \ldots, \qquad x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m = 0$$

Mas $\{a_1, ..., a_m\}$ é linearmente independente sobre K; como $x_{ji} \in K$, decorre que

$$x_{11} = 0,\ x_{12} = 0,\ \ldots,\ x_{1m} = 0,\ \ldots,\ x_{n1} = 0,\ x_{n2} = 0,\ \ldots,\ x_{nm} = 0$$

Dessa forma, $\{a_i v_j\}$ é linearmente independente sobre K, provando a afirmação.

Problemas Complementares

Espaços vetoriais

4.71 Suponha que u e v pertençam a um espaço vetorial V. Simplifique as expressões dadas.

(a) $E_1 = 4(5u - 6v) + 2(3u + v)$, (c) $E_3 = 6(3u + 2v) + 5u - 7v$,

(b) $E_2 = 5(2u - 3v) + 4(7v + 8)$, (d) $E_4 = 3(5u + 2/v)$.

4.72 Seja V o conjunto dos pares ordenados (a, b) de números reais com adição e multiplicação por escalar definidas por

$$(a, b) + (c, d) = (a + c, \quad b + d) \qquad \text{e} \qquad k(a, b) = (ka, 0)$$

Mostre que V satisfaz todos os axiomas de espaço vetorial exceto $[M_4]$, ou seja, que $1u = u$. Logo, $[M_4]$ não é uma consequência dos demais axiomas.

4.73 Mostre que o Axioma $[A_4]$ de um espaço vetorial V (ou seja, que $u + v = v + u$) pode ser deduzido dos demais axiomas de V.

4.74 Seja V o conjunto de todos os pares ordenados (a, b) de números reais. Mostre que V não é um espaço vetorial sobre **R** com adição e multiplicação por escalar definidas por

(i) $(a,b)+(c,d)=(a+d,\ b+c)$ e $k(a,b)=(ka,kb)$,
(ii) $(a,b)+(c,d)=(a+c,\ b+d)$ e $k(a,b)=(a,b)$,
(iii) $(a,b)+(c,d)=(0,0)$ e $k(a,b)=(ka,kb)$,
(iv) $(a,b)+(c,d)=(ac,bd)$ e $k(a,b)=(ka,kb)$.

4.75 Seja V o conjunto das sequências infinitas $(a_1, a_2, ...)$ de um corpo K. Mostre que V é um espaço vetorial sobre K com adição e multiplicação por escalar definidas por

$$(a_1,a_2,\ldots)+(b_1,b_2,\ldots)=(a_1+b_1,\ a_2+b_2,\ \ldots) \quad \text{e} \quad k(a_1,a_2,\ldots)=(ka_1,ka_2,\ldots)$$

4.76 Sejam U e W espaços vetoriais sobre um corpo K. Seja V o conjunto dos pares ordenados (u, w), com $u \in U$ e $w \in W$. Mostre que V é um espaço vetorial sobre K com adição e multiplicação por escalar definidas por

$$(u,w)+(u',w')=(u+u',\ w+w') \quad \text{e} \quad k(u,w)=(ku,kw)$$

(Esse espaço V é denominado *produto cartesiano* de U e W.)

Subespaços

4.77 Decida se W é, ou não, um subespaço de $\mathbf{R}^3$, se W consistir de todos os vetores (a, b, c) de $\mathbf{R}^3$ tais que (a) $a = 3b$, (b) $a \le b \le c$, (c) $ab = 0$, (d) $a+b+c=0$, (e) $b=a^2$, (*f*) $a=2b=3c$.

4.78 Seja V o espaço vetorial das matrizes quadradas de ordem n sobre um corpo K. Mostre que W é um subespaço de V se W consistir em todas as matrizes $A=[a_{ij}]$ que são
(a) simétricas ($A^T = A$ ou $a_{ij}=a_{ji}$), (b) triangulares (superiores), (c) diagonais, (d) escalares .

4.79 Seja $AX = B$ um sistema não homogêneo de equações lineares com n incógnitas, ou seja, $B \neq 0$. Mostre que o conjunto solução não é um subespaço vetorial de K^n.

4.80 Suponha que U e W sejam subespaços de V tais que $U \cup W$ seja um subespaço. Mostre que $U \subseteq W$ ou $W \subseteq U$.

4.81 Seja V o espaço vetorial de todas as funções do corpo real $\mathbf{R}$ em $\mathbf{R}$. Mostre que W é um subespaço de V se W consiste em todas as (a) funções limitadas, (b) funções pares. [Lembre que $f\colon \mathbf{R} \to \mathbf{R}$ é *limitada* se existir $M \in \mathbf{R}$ tal que $|f(x)| \le M$, para cada $x \in \mathbf{R}$ e que f é *par* se $f(-x)=f(x), \forall x \in \mathbf{R}$.]

4.82 Seja V o espaço vetorial das sequências infinitas $(a_1, a_2, ...)$ de um corpo K (Problema 4.75). Mostre que W é um subespaço de V se W consiste em todas as sequências (a) com primeiro elemento $a_1 = 0$, (b) com um número finito de elementos não nulos.

Combinações lineares, espaços gerados

4.83 Considere os vetores $u = (1, 2, 3)$ e $v = (2, 3, 1)$ de $\mathbf{R}^3$.

(a) Escreva $w = (1, 3, 8)$ como uma combinação linear de u e v.

(b) Escreva $w = (2, 4, 5)$ como uma combinação linear de u e v.

(c) Encontre k de tal modo que $w = (1, k, 4)$ seja uma combinação linear de u e v.

(d) Encontre condições sobre a, b, c tais que $w = (a, b, c)$ seja uma combinação linear de u e v.

4.84 Escreva o polinômio $f(t)=at^2+bt+c$ como uma combinação linear dos polinômios $p_1=(t-1)^2$, $p_2=t-1$, $p_3=1$. [Assim, p_1, p_2, p_3 geram o espaço $\mathbf{P}_2(t)$ dos polinômios de grau ≤ 2.]

4.85 Encontre um vetor de $\mathbf{R}^3$ que gere a interseção de U e W, sendo U o plano xy, ou seja, $U = \{(a, b, 0)\}$, e W o espaço gerado pelos vetores (1, 1, 1) e (1, 2, 3).

4.86 Demonstre que ger(S) é a interseção de todos os subespaços de V que contém S.

4.87 Mostre que ger(S) = ger($S \cup \{0\}$). Ou seja, acrescentando ou retirando o vetor nulo de um conjunto, não modificamos o espaço gerado pelo conjunto.

4.88 Mostre que (a) se $S \subseteq T$ então $\text{ger}(S) \subseteq \text{ger}(T)$. (b) $\text{ger}[\text{ger}(S)] = \text{ger}(S)$

Dependência e independência linear

4.89 Decida se os vetores dados são linearmente dependentes ou independentes.

(a) $(1, 2, -3, 1), (3, 7, 1, -2), (1, 3, 7, -4)$; (b) $(1, 3, 1, -2), (2, 5, -1, 3), (1, 3, 7, -2)$.

4.90 Decida se os polinômios u, v, w de $\mathbf{P}(t)$ são linearmente dependentes ou independentes.

(a) $u = t^3 - 4t^2 + 3t + 3, \quad v = t^3 + 2t^2 + 4t - 1, \quad w = 2t^3 - t^2 - 3t + 5$;

(b) $u = t^3 - 5t^2 - 2t + 3, \quad v = t^3 - 4t^2 - 3t + 4, \quad w = 2t^3 - 17t^2 - 7t + 9$.

4.91 Mostre que as funções f, g, h dadas são linearmente independentes.

(a) $f(t) = e^t, \quad g(t) = \text{sen}\, t, \quad h(t) = t^2$; (b) $f(t) = e^t, \quad g(t) = e^{2t}, \quad h(t) = t$.

4.92 Mostre que $u = (a, b)$ e $v = (c, d)$ de K^2 são linearmente dependentes se, e só se, $ad - bc = 0$.

4.93 Sejam u, v, w vetores linearmente independentes. Mostre que os conjuntos S dados são linearmente independentes.

(a) $S = \{u + v - 2w, \ u - v - w, \ u + w\}$; (b) $S = \{u + v - 3w, \ u + 3v - w, \ v + w\}$.

4.94 Suponha que $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ seja um subconjunto linearmente independente de V. Mostre que

$$\text{ger}(u_i) \cap \text{ger}(w_j) = \{0\}$$

4.95 Suponha que $v_1, v_2, \ldots, v_n$ sejam linearmente independentes. Mostre que S é linearmente independente nos casos dados.

(a) $S = \{a_1 v_1, a_2 v_2, \ldots, a_n v_n\}$ e cada $a_i \neq 0$.

(b) $S = \{v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n\}$ e $w = \sum_i b_i v_i$ com $b_k \neq 0$.

4.96 Suponha que $(a_{11}, \ldots, a_{1n}), \ (a_{21}, \ldots, a_{2n}), \ \ldots, \ (a_{m1}, \ldots, a_{mn})$ sejam vetores linearmente independente de K^n e suponha que $v_1, v_2, \ldots, v_n$ sejam vetores linearmente independentes de um espaço vetorial V sobre K. Mostre que os vetores dados também são linearmente independentes.

$$w_1 = a_{11} v_1 + \cdots + a_{1n} v_n, \qquad w_2 = a_{21} v_1 + \cdots + a_{2n} v_n, \qquad \ldots, \qquad w_m = a_{m1} v_1 + \cdots + a_{mn} v_n$$

Base e dimensão

4.97 Encontre um subconjunto de u_1, u_2, u_3, u_4 que seja uma base de $W = \text{ger}(u_i)$ em $\mathbf{R}^5$, nos casos seguintes.

(a) $u_1 = (1, 1, 1, 2, 3), \quad u_2 = (1, 2, -1, -2, 1), \quad u_3 = (3, 5, -1, -2, 5), \quad u_4 = (1, 2, 1, -1, 4)$

(b) $u_1 = (1, -2, 1, 3, -1), \quad u_2 = (-2, 4, -2, -6, 2), \quad u_3 = (1, -3, 1, 2, 1), \quad u_4 = (3, -7, 3, 8, -1)$

(c) $u_1 = (1, 0, 1, 0, 1), \quad u_2 = (1, 1, 2, 1, 0), \quad u_3 = (2, 1, 3, 1, 1), \quad u_4 = (1, 2, 1, 1, 1)$

(d) $u_1 = (1, 0, 1, 1, 1), \quad u_2 = (2, 1, 2, 0, 1), \quad u_3 = (1, 1, 2, 3, 4), \quad u_4 = (4, 2, 5, 4, 6)$

4.98 Considere os subespaços $U = \{(a, b, c, d) : b - 2c + d = 0\}$ e $W = \{(a, b, c, d) : a = d, b = 2c\}$ de $\mathbf{R}^4$. Encontre uma base e a dimensão de (a) U, (b) W, (c) $U \cap W$.

4.99 Encontre uma base e a dimensão do espaço solução W de cada um dos sistemas homogêneos dados.

$$\text{(a)} \quad \begin{aligned} x + 2y - 2z + 2s - \ t &= 0 \\ x + 2y - \ z + 3s - 2t &= 0 \\ 2x + 4y - 7z + \ s + \ t &= 0 \end{aligned} \qquad \text{(b)} \quad \begin{aligned} x + 2y - \ z + 3s - 4t &= 0 \\ 2x + 4y - 2z - \ s + 5t &= 0 \\ 2x + 4y - 2z + 4s - 2t &= 0 \end{aligned}$$

4.100 Encontre um sistema homogêneo cujo espaço solução seja gerado pelos conjuntos de três vetores dados.

(a) $(1, -2, 0, 3, -1), (2, -3, 2, 5, -3), (1, -2, 1, 2, -2)$;

(b) $(1, 1, 2, 1, 1), (1, 2, 1, 4, 3), (3, 5, 4, 9, 7)$.

4.101 Decida se os conjuntos dados são bases do espaço vetorial $\mathbf{P}_n(t)$.

(a) $\{1, \quad 1+t, \quad 1+t+t^2, \quad 1+t+t^2+t^3, \quad \ldots, \quad 1+t+t^2+\cdots+t^{n-1}+t^n\}$;

(b) $\{1+t, \quad t+t^2, \quad t^2+t^3, \quad \ldots, \quad t^{n-2}+t^{n-1}, \quad t^{n-1}+t^n\}$.

4.102 Encontre uma base e a dimensão do subespaço W de $\mathbf{P}(t)$ gerado por

(a) $u = t^3+2t^2-2t+1, \quad v = t^3+3t^2-3t+4, \quad w = 2t^3+t^2-7t+11,$

(b) $u = t^3+t^2-3t+2, \quad v = 2t^3+t^2+t-4, \quad w = 4t^3+3t^2-5t+2.$

4.103 Encontre uma base e a dimensão do subespaço W de $\mathbf{M}_{2,2}$ gerado por

$$A = \begin{bmatrix} 1 & -5 \\ -4 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ -1 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -4 \\ -5 & 7 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & -7 \\ -5 & 1 \end{bmatrix}$$

Posto de uma matriz, espaços linha e coluna

4.104 Encontre o posto de cada uma das matrizes dadas.

$$\text{(a)} \begin{bmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} 1 & 2 & -3 & -2 \\ 1 & 3 & -2 & 0 \\ 3 & 8 & -7 & -2 \\ 2 & 1 & -9 & -10 \end{bmatrix}, \quad \text{(c)} \begin{bmatrix} 1 & 1 & 2 \\ 4 & 5 & 5 \\ 5 & 8 & 1 \\ -1 & -2 & 2 \end{bmatrix}$$

4.105 Para $k = 1, 2, ..., 5$, encontre o número n_k de subconjuntos linearmente independentes constituídos de k colunas de cada uma das matrizes dadas.

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 1 & 0 & 2 & 3 \\ 1 & 2 & 0 & 2 & 5 \\ 1 & 3 & 0 & 2 & 7 \end{bmatrix}, \quad \text{(b)} \quad B = \begin{bmatrix} 1 & 2 & 1 & 0 & 2 \\ 1 & 2 & 3 & 0 & 4 \\ 1 & 1 & 5 & 0 & 6 \end{bmatrix}$$

4.106 Sejam (a) $A = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 6 \\ 2 & 4 & 3 & 8 & 3 & 15 \\ 1 & 2 & 2 & 5 & 3 & 11 \\ 4 & 8 & 6 & 16 & 7 & 32 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 2 & 1 & 2 & 1 \\ 2 & 4 & 5 & 4 & 5 & 5 \\ 1 & 2 & 3 & 4 & 4 & 6 \\ 3 & 6 & 7 & 7 & 9 & 10 \end{bmatrix}$

e denote por $C_1, ..., C_6$ suas colunas. Para cada uma das duas matrizes,

(i) encontre a forma canônica por linhas M;

(ii) encontre as colunas que são combinações lineares das colunas que as antecedem;

(iii) encontre as colunas (excetuando C_6) que formam uma base do espaço coluna;

(iv) escreva C_6 como uma combinação linear dos vetores da base encontrada em (iii).

4.107 Decida se as matrizes dadas têm o mesmo espaço linha.

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 3 & -4 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 3 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 & 3 \\ 2 & -1 & 10 \\ 3 & -5 & 1 \end{bmatrix}$$

4.108 Encontre os subespaços de $\mathbf{R}^3$ idênticos dentre os dados.

$$U_1 = \text{ger}\,[(1,1,-1),\ (2,3,-1),\ (3,1,-5)], \qquad U_2 = \text{ger}\,[(1,-1,-3),\ (3,-2,-8),\ (2,1,-3)]$$
$$U_3 = \text{ger}\,[(1,1,1),\ (1,-1,3),\ (3,-1,7)]$$

4.109 Encontre os subespaços de $\mathbf{R}^4$ idênticos dentre os dados.

$$U_1 = \text{ger}\,[(1,2,1,4),\ (2,4,1,5),\ (3,6,2,9)], \qquad U_2 = \text{ger}\,[(1,2,1,2),\ (2,4,1,3)],$$
$$U_3 = \text{ger}\,[(1,2,3,10),\ (2,4,3,11)]$$

4.110 Encontre uma base do (i) espaço linha e do (ii) espaço coluna de cada matriz M dada.

(a) $M = \begin{bmatrix} 0 & 0 & 3 & 1 & 4 \\ 1 & 3 & 1 & 2 & 1 \\ 3 & 9 & 4 & 5 & 2 \\ 4 & 12 & 8 & 8 & 7 \end{bmatrix}$, (b) $M = \begin{bmatrix} 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 3 \\ 3 & 6 & 5 & 2 & 7 \\ 2 & 4 & 1 & -1 & 0 \end{bmatrix}$.

4.111 Mostre que se uma linha qualquer for suprimida de uma matriz em forma escalonada (respectivamente, forma canônica por linhas), então a matriz resultante ainda está em forma escalonada (respectivamente, forma canônica por linhas).

4.112 Sejam A e B matrizes $m \times n$ quaisquer. Mostre que $\text{pos}(A+B) \leq \text{pos}(A) + \text{pos}(B)$.

4.113 Seja $r = \text{pos}(A + B)$. Encontre matrizes quadradas A e B de ordem 2 tais que

(a) $r < \text{pos}(A), \text{pos}(B)$; (b) $r = \text{pos}(A) = \text{pos}(B)$; (c) $r > \text{pos}(A), \text{pos}(B)$.

Somas, somas diretas, interseções

4.114 Suponha que U e W sejam subespaços bidimensionais de K^3. Mostre que $U \cap W \neq \{0\}$.

4.115 Suponha que U e W sejam subespaços de V tais que $\dim U = 4$, $\dim W = 5$ e $\dim V = 7$. Encontre as dimensões possíveis de $U \cap W$.

4.116 Sejam U e W subespaços de $\mathbf{R}^3$ tais que $\dim U = 1$, $\dim W = 2$ e $U \not\subseteq W$. Mostre que $\mathbf{R}^3 = U \oplus W$.

4.117 Considere os subespaços de $\mathbf{R}^5$ dados.

$$U = \text{ger}[(1, -1, -1, -2, 0), (1, -2, -2, 0, -3), (1, -1, -2, -2, 1)]$$

$$W = \text{ger}[(1, -2, -3, 0, -2), (1, -1, -3, 2, -4), (1, -1, -2, 2, -5)]$$

(a) Encontre dois sistemas homogêneos cujos espaços solução sejam U e W, respectivamente.

(b) Encontre uma base e a dimensão de $U \cap W$.

4.118 Sejam U_1, U_2, U_3 os subespaços de $\mathbf{R}^3$ dados.

$$U_1 = \{(a, b, c) : a = c\}, \qquad U_2 = \{(a, b, c) : a + b + c = 0\}, \qquad U_3 = \{(0, 0, c)\}$$

Mostre que (a) $\mathbf{R}^3 = U_1 + U_2$, (b) $\mathbf{R}^3 = U_2 + U_3$, (c) $\mathbf{R}^3 = U_1 + U_3$. Quais dessas somas são diretas?

4.119 Suponha que U, W_1, W_2 sejam subespaços de um espaço vetorial V. Mostre que

$$(U \cap W_1) + (U \cap W_2) \subseteq U \cap (W_1 + W_2)$$

Encontre subespaços de $\mathbf{R}^2$ para os quais a inclusão é própria.

4.120 Suponha que $W_1, W_2, ..., W_r$ sejam subespaços de um espaço vetorial V. Mostre que

(a) $\text{ger}(W_1, W_2, \ldots, W_r) = W_1 + W_2 + \cdots + W_r$.

(b) Se S_i gera W_i, com $i = 1, ..., r$, então $S_1 \cup S_2 \cup \cdots \cup S_r$ gera $W_1 + W_2 + \cdots + W_r$.

4.121 Suponha que $V = U \oplus W$. Mostre que $\dim V = \dim U + \dim W$.

4.122 Sejam S e T subconjuntos (não necessariamente subespaços) arbitrários não vazios de um espaço vetorial V e seja r um escalar. A soma $S + T$ e o múltiplo rS são definidos por

$$S + T = \{u + v : u \in S, \ v \in T\}, \qquad rS = \{ru : u \in S\}$$

[Também escrevemos $w + S$ em vez de $\{w\} + S$.] Sejam

$$S = \{(1, 2), (2, 3)\}, T = \{(1, 4), (1, 5), (2, 5)\}, w = (1, 1), r = 3$$

Encontre (a) $S + T$, (b) $w + S$, (c) rT, (d) $rS + rT$, (*f*) $r(S + T)$.

4.123 Mostre que as operações $S + T$ e rS do problema precedente satisfazem as afirmações dadas.

(a) Comutatividade: $S + T = T + S$.

(b) Associatividade: $(S_1 + S_2) + S_3 = S_1 + (S_2 + S_3)$.

(c) Distributividade: $r(S + T) = rS + rT$.

(d) $S + \{0\} = \{0\} + S = S$ e $S + V = V + S = V$.

4.124 Seja V o espaço vetorial das matrizes quadradas de ordem n. Seja U o subespaço das matrizes triangulares superiores e seja W o subespaço das matrizes triangulares inferiores. Encontre (a) $U \cap W$, (b) $U + W$.

4.125 Seja V o produto cartesiano dos espaços vetoriais U e W sobre K (ver Problema 4.76). Sejam

$$\hat{U} = \{(u, 0) : u \in U\} \qquad \text{e} \qquad \hat{W} = \{(0, w) : w \in W\}$$

Mostre que (a) $\hat{U}$ e $\hat{W}$ são subespaços de V, (b) $V = \hat{U} \oplus \hat{W}$

4.126 Suponha que $V = U + W$. Seja $\hat{V}$ o produto cartesiano de U e W. Mostre que V é isomorfo a $\hat{V}$ pela correspondência $v = u + w \leftrightarrow (u, w)$.

4.127 Use indução para provar (a) o Teorema 4.22, (b) o Teorema 4.23.

Coordenadas

4.128 Os vetores $u_1 = (1, -2)$ e $u_2 = (4, -7)$ formam uma base S de $\mathbf{R}^2$. Encontre o vetor de coordenadas $[v]$ de v em relação a S nos casos (a) $v = (5, 3)$, (b) $v = (a, b)$.

4.129 Os vetores $u_1 = (1, 2, 0)$, $u_2 = (1, 3, 2)$, $u_3 = (0, 1, 3)$ formam uma base S de $\mathbf{R}^3$. Encontre o vetor de coordenadas $[v]$ de v em relação a S nos casos (a) $v = (2, 7, -4)$, (b) $v = (a, b, c)$.

4.130 $S = \{t^3 + t^2, \quad t^2 + t, \quad t + 1, \quad 1\}$ é uma base de $\mathbf{P}_3(t)$. Encontre o vetor de coordenadas $[v]$ de v em relação a S nos casos (a) $v = 2t^3 + t^2 - 4t + 2$, (b) $v = at^3 + bt^2 + ct + d$.

4.131 Seja $V = \mathbf{M}_{2,2}$. Encontre o vetor de coordenadas $[A]$ de A em relação a S, sendo

$$S = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\} \quad \text{e} \quad \text{(a)} \quad A = \begin{bmatrix} 3 & -5 \\ 6 & 7 \end{bmatrix}, \quad \text{(b)} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

4.132 Encontre a dimensão e uma base do subespaço W de $\mathbf{P}_3(t)$ gerado por

$$u = t^3 + 2t^2 - 3t + 4, \qquad v = 2t^3 + 5t^2 - 4t + 7, \qquad w = t^3 + 4t^2 + t + 2$$

4.133 Encontre a dimensão e uma base do subespaço W de $\mathbf{M} = \mathbf{M}_{2,3}$ gerado por

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 4 & 3 \\ 7 & 5 & 6 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 6 \end{bmatrix}$$

Problemas variados

4.134 Responda com verdadeiro ou falso. Se falso, justifique com um contraexemplo.

(a) Se u_1, u_2, u_3 geram V, então dim $V = 3$.

(b) Se A é uma matriz 4×8, então quaisquer seis colunas são linearmente dependentes.

(c) Se u_1, u_2, u_3 são linearmente independentes, então u_1, u_2, u_3, w são linearmente dependentes.

(d) Se u_1, u_2, u_3, u_4 são linearmente independentes, então $V \geq 4$.

(e) Se u_1, u_2, u_3 geram V, então w, u_1, u_2, u_3 geram V.

(f) Se u_1, u_2, u_3, u_4 são linearmente independentes, então u_1, u_2, u_3 são linearmente independentes.

4.135 Responda com verdadeiro ou falso. Se falso, justifique com um contraexemplo.

(a) Se uma coluna qualquer for suprimida de uma matriz em forma escalonada, então a matriz resultante ainda está em forma escalonada.

(b) Se uma coluna qualquer for suprimida de uma matriz em forma canônica por linhas, então a matriz resultante ainda está em forma canônica por linhas.

(c) Se uma coluna qualquer sem pivô for suprimida de uma matriz em forma canônica por linhas, então a matriz resultante ainda está em forma canônica por linhas.

4.136 Encontre a dimensão do espaço vetorial W das matrizes quadradas de ordem n dadas.

(a) matrizes simétricas, (b) matrizes antissimétricas,

(c) matrizes diagonais, (d) matrizes escalares.

4.137 Sejam $t_1, t_2, ..., t_n$ símbolos e K um corpo arbitrário. Seja V o conjunto seguinte de expressões, em que $a_i \in K$:

$$a_1t_1 + a_2t_2 + \cdots + a_nt_n$$

Defina a soma e multiplicação por escalar em V por

$$(a_1t_1 + \cdots + a_nt_n) + (b_1t_1 + \cdots + b_nt_n) = (a_1 + b_1)t_1 + \cdots + (a_nb_{nm})t_n$$
$$k(a_1t_1 + a_2t_2 + \cdots + a_nt_n) = ka_1t_1 + ka_2t_2 + \cdots + ka_nt_n$$

Mostre que, com essas operações, V é um espaço vetorial sobre K. Também mostre que $\{t_1, t_2, ..., t_n\}$ é uma base de V, em que

$$t_j = 0t_1 + \cdots + 0t_{j-1} + 1t_j + 0t_{j+1} + \cdots + 0t_n$$

Respostas dos Problemas Complementares

[Algumas respostas, tais como bases, não precisam ser únicas.]

4.71 (a) $E_1 = 26u - 22v$; (b) A soma $7v + 8$ não está definida, portanto, E_2 não está definida;
(c) $E_3 = 23u + 5v$; (d) A divisão por v não está definida, portanto, E_4 não está definida.

4.77 (a) É; (b) Não é; por exemplo, $(1, 2, 3) \in W$, mas $-2(1, 2, 3) \notin W$;
(c) Não é; por exemplo, $(1, 0, 0), (0, 1, 0) \in W$, mas não sua soma; (d) É;
(e) Não é; por exemplo, $(1, 1, 1) \in W$, mas $2(1, 1, 1) \notin W$; (f) É.

4.79 O vetor nulo não é uma solução.

4.83 (a) $w = 3u_1 - u_2$, (b) Impossível, (c) $k = \frac{11}{5}$, (d) $7a - 5b + c = 0$

4.84 Usando $f = xp_1 + yp_2 + zp_3$, obtemos $x = a,\ y = 2a + b,\ z = a + b + c$.

4.85 $v = (2, 1, 0)$

4.89 (a) Dependentes, (b) Independentes

4.90 (a) Independentes, (b) Dependentes

4.97 (a) u_1, u_2, u_4; (b) u_1, u_2, u_3; (c) u_1, u_2, u_4; (d) u_1, u_2, u_3

4.98 (a) $\dim U = 3$, (b) $\dim W = 2$, (c) $\dim(U \cap W) = 1$

4.99 (a) Base: $\{(2, -1, 0, 0, 0),\ (4, 0, 1, -1, 0),\ (3, 0, 1, 0, 1)\}$; $\dim W = 3$;
(b) Base: $\{(2, -1, 0, 0, 0),\ (1, 0, 1, 0, 0)\}$; $\dim W = 2$

4.100 (a) $5x + y - z - s = 0, \quad x + y - z - t = 0$;
(b) $3x - y - z = 0, \quad 2x - 3y + s = 0, \quad x - 2y + t = 0$

4.101 (a) É, (b) Não é, porque $\dim \mathbf{P}_n(t) = n + 1$, mas o conjunto contém somente n elementos.

4.102 (a) $\dim W = 2$, (b) $\dim W = 3$

4.103 $\dim W = 2$

4.104 (a) 3, (b) 2, (c) 3

4.105 (a) $n_1 = 4, \quad n_2 = 5, \quad n_3 = n_4 = n_5 = 0$; (b) $n_1 = 4, \quad n_2 = 6, \quad n_3 = 3, \quad n_4 = n_5 = 0$

4.106 (a) (i) $M = [1, 2, 0, 1, 0, 3; \quad 0, 0, 1, 2, 0, 1; \quad 0, 0, 0, 0, 1, 2; \quad 0]$;
(ii) C_2, C_4, C_6; (iii) C_1, C_3, C_5; (iv) $C_6 = 3C_1 + C_3 + 2C_5$.
(b) (i) $M = [1, 2, 0, 0, 3, 1; \quad 0, 0, 1, 0, -1, -1; \quad 0, 0, 0, 1, 1, 2; \quad 0]$;
(ii) C_2, C_5, C_6; (iii) C_1, C_3, C_4; (iv) $C_6 = C_1 - C_3 + 2C_4$

4.107 A e C são equivalentes por linhas a $\begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 4 \end{bmatrix}$, mas B não é.

4.108 U_1 e U_2 são equivalentes por linhas a $\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}$, mas U_3 não é.

4.109 U_1 e U_3 são equivalentes por linhas a $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$, mas U_2 não é.

4.110 (a) (i) $(1, 3, 1, 2, 1), \ (0, 0, 1, -1, -1), \ (0, 0, 0, 4, 7)$; (ii) C_1, C_3, C_4;
(b) (i) $(1, 2, 1, 0, 1), \ (0, 0, 1, 1, 2)$; (ii) C_1, C_3

4.113 (a) $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix}$; (b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$;
(c) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

4.115 $\dim(U \cap W) = 2$, 3 ou 4

4.117 (a) (i) $\begin{matrix} 3x + 4y - z - t = 0 \\ 4x + 2y + s = 0 \end{matrix}$ (ii) $\begin{matrix} 4x + 2y - s = 0 \\ 9x + 2y + z + t = 0 \end{matrix}$;
(b) Base: $\{(1, -2, -5, 0, 0), \quad (0, 0, 1, 0, -1)\}$; $\dim(U \cap W) = 2$

4.118 A soma é direta em (b) e (c).

4.119 Em $\mathbf{R}^2$, basta tomar U, V, W, respectivamente, a reta $y = x$, o eixo x e o eixo y.

4.122 (a) $\{(2, 6), \quad (2, 7), \quad (3, 7), \quad (3, 8), \quad (4, 8)\}$; (b) $\{(2, 3), \quad (3, 4)\}$;
(c) $\{(3, 6), \quad (6, 9)\}$; (d) $\{(3, 12), \quad (3, 15), \quad (6, 15)\}$;
(e e f) $\{(6, 18), \quad (6, 21), \quad (9, 21), \quad (9, 24), \quad (12, 24)\}$

4.124 (a) Matrizes diagonais, (b) V

4.128 (a) $[-41, 11]$, (b) $[-7a - 4b, \quad 2a + b]$

4.129 (a) $[-11, 13, -10]$, (b) $[c - 3b + 7a, \quad -c + 3b - 6a, \quad c - 2b + 4a]$

4.130 (a) $[2, -1, -2, 2]$, (b) $[a, \quad b - c, \quad c - b + a, \quad d - c + b - a]$

4.131 (a) $[7, -1, -13, 10]$, (b) $[d, \quad c - d, \quad b + c - 2d, \quad a - b - 2c + 2d]$

4.132 dim $W = 2$; base $\{t^3 + 2t^2 - 3t + 4, \quad t^2 + 2t - 1\}$.

4.133 dim $W = 2$; base $\{[1, 2, 1, 3, 1, 2], \quad [0, 0, 1, 1, 3, 2]\}$.

4.134 (a) Falsa; $(1, 1),\ (1, 2),\ (2, 1)$ geram $\mathbf{R}^2$; (b) Verdadeira;
(c) Falsa; $(1, 0, 0, 0),\ (0, 1, 0, 0),\ (0, 0, 1, 0),\ w = (0, 0, 0, 1)$;
(d) Verdadeira; (e) Verdadeira; (f) Verdadeira

4.135 (a) Verdadeira; (b) Falsa; por exemplo, suprimindo C_2 de $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$; (c) Verdadeira

4.136 (a) $\frac{1}{2}n(n+1)$, (b) $\frac{1}{2}n(n-1)$, (c) n, (d) 1

Transformações Lineares

5.1 INTRODUÇÃO

O assunto principal da Álgebra Linear é o estudo das transformações lineares e suas representações por meio de matrizes. Neste capítulo apresentamos essas transformações lineares e, no próximo, mostramos como elas podem ser representadas por matrizes. Inicialmente, contudo, apresentamos um estudo de aplicações em geral.

5.2 APLICAÇÕES, FUNÇÕES

Sejam A e B dois conjuntos não vazios quaisquer. Suponha que a cada elemento $a \in A$ esteja associado um único elemento de B, denominado *imagem de a*. A coleção f dessas associações é denominada *aplicação* de A em B e é denotada por

$$f : A \to B$$

Dizemos que o conjunto A é o *domínio* da aplicação e que B é o *contradomínio*. Escrevemos $f(a)$, que é lido "efe de a", para o único elemento de B que f associa ao elemento $a \in A$.

Podemos também interpretar uma aplicação $f : A \to B$ como um computador que, para cada valor inicial $a \in A$, produz um valor final $f(a) \in B$.

OBSERVAÇÃO O termo *função* é um sinônimo da palavra *aplicação*, embora alguns textos reservem a palavra "função" para aplicações de valores reais ou complexos.

Considere uma aplicação $f : A \to B$. Se A' for um subconjunto qualquer de A, então $f(A')$ denota o conjunto das imagens de todos elementos de A'. Se B' for um subconjunto qualquer de B, então $f^{-1}(B')$ denota o conjunto de todos os elementos de A cujas imagens estão em B. Ou seja,

$$f(A') = \{f(a) : a \in A'\} \qquad \text{e} \qquad f^{-1}(B') = \{a \in A : f(a) \in B'\}$$

Dizemos que $f(A')$ é a *imagem* de A' e que $f^{-1}(B')$ é a *imagem inversa*, ou *pré-imagem*, de B'. Em particular, dizemos que o conjunto de todas as imagens de f, ou seja, $f(A)$, é a *imagem* de f.

A cada aplicação $f : A \to B$ corresponde o subconjunto de $A \times B$ dado por $\{(a, f(a)) : a \in A\}$, que é denominado *gráfico* de f. Duas aplicações $f : A \to B$ e $g : A \to B$ são definidas como sendo *iguais*, e escrevemos $f = g$, se $f(a) = g(a)$, para cada $a \in A$, ou seja, se possuírem o mesmo gráfico. Assim, não há distinção entre uma função e seu gráfico. A negação de $f = g$ é denotada por $f \neq g$ e significa

$$\boxed{\text{Existe algum } a \in A \text{ para o qual } f(a) \neq g(a).}$$

Às vezes se utiliza a seta especial $\mapsto$ para denotar a imagem de um elemento arbitrário $x \in A$ por uma aplicação $f : A \to B$, escrevendo

$$x \mapsto f(x)$$

como no exemplo seguinte.

Exemplo 5.1

(a) Seja $f : \mathbf{R} \to \mathbf{R}$ a função que a cada número real x associa seu quadrado x^2. Podemos escrever essa função como segue.

$$f(x) = x^2 \qquad \text{ou} \qquad x \mapsto x^2$$

Aqui, a imagem de –3 é 9, portanto podemos escrever $f(-3) = 9$. Contudo, $f^{-1}(9) = \{3, -3\}$. Também $f(\mathbf{R}) = [0, \infty) = \{x : x \geq 0\}$ é a imagem de f.

(b) Sejam $A = \{a, b, c, d\}$ e $B = \{x, y, z, t\}$. Uma aplicação $f : A \to B$ pode ser definida por

$$f(a) = y,\ f(b) = x,\ f(c) = z,\ f(d) = y \qquad \text{ou} \qquad f = \{(a, y),\ (b, x),\ (c, z),\ (d, y)\}$$

A primeira definição é explícita e a segunda é dada pelo gráfico. Temos

$$f(\{a, b, d\}) = \{f(a), f(b), f(d)\} = \{y, x, y\} = \{x, y\}$$

Além disso, a imagem de f é $f(A) = \{x, y, z\}$.

Exemplo 5.2 Seja V o espaço vetorial dos polinômios sobre $\mathbf{R}$ e seja $p(t) = 3t^2 - 5t + 2$.

(a) A derivada define uma aplicação $\mathbf{D} : V \to V$, em que, dado um polinômio $f(t)$ qualquer, temos $\mathbf{D}(f) = df/dt$. Assim,

$$\mathbf{D}(p) = \mathbf{D}(3t^2 - 5t + 2) = 6t - 5$$

(b) A integral, digamos, de 0 a 1, define uma função $\mathbf{J} : V \to \mathbf{R}$. Ou seja, dado um polinômio $f(t)$ qualquer,

$$\mathbf{J}(f) = \int_0^1 f(t)\, dt, \qquad \text{e, portanto,} \qquad \mathbf{J}(p) = \int_0^1 (3t^2 - 5t + 2) = \tfrac{1}{2}.$$

Observe que a aplicação de (b) é do espaço vetorial V no corpo de escalares $\mathbf{R}$, enquanto a aplicação de (a) é do espaço vetorial V nele mesmo.

Aplicações matriciais

Seja A uma matriz $m \times n$ sobre K. Então A determina uma aplicação $F_A : K^n \to K^m$ por meio de

$$F_A(u) = Au$$

em que os vetores de K^n e K^m estão escritos como colunas. Por exemplo, suponha que

$$A = \begin{bmatrix} 1 & -4 & 5 \\ 2 & 3 & -6 \end{bmatrix} \qquad \text{e} \qquad u = \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix}$$

então

$$F_A(u) = Au = \begin{bmatrix} 1 & -4 & 5 \\ 2 & 3 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix} = \begin{bmatrix} -36 \\ 41 \end{bmatrix}$$

OBSERVAÇÃO Por conveniência notacional, muitas vezes escrevemos a aplicação F_A com a mesma letra A utilizada para a matriz.

Composição de aplicações

Considere duas aplicações $f : A \to B$ e $g : B \to C$, ilustradas a seguir.

$$A \xrightarrow{f} B \xrightarrow{g} C$$

A *composição* de *f* e *g*, denotada por $g \circ f$, é a aplicação $g \circ f : A \to C$ definida por

$$(g \circ f)(a) \equiv g(f(a))$$

Isto é, primeiro aplicamos *f* a $a \in A$ e depois aplicamos *g* a $f(a) \in B$ para obter $g(f(a)) \in C$. Interpretando *f* e *g* como "computadores", a composição significa que primeiro utilizamos em *f* o dado inicial $a \in A$ para obter a resposta $f(a) \in B$ e depois utilizamos em *g* o dado inicial *f*(*a*) para obter a resposta $g(f(a)) \in C$.

Nosso primeiro teorema diz que a composição de aplicações é associativa.

Teorema 5.1 Sejam $f : A \to B$, $g : B \to C$, $h : C \to D$. Então

$$h \circ (g \circ f) = (h \circ g) \circ f$$

Demonstremos esse teorema. Seja $a \in A$. Então

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a)))$$
$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a)))$$

Assim, $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$, para cada $a \in A$, portanto, $h \circ (g \circ f) = (h \circ g) \circ f$.

Aplicações injetoras e sobrejetoras

Introduzimos formalmente alguns tipos especiais de aplicações.

DEFINIÇÃO Dizemos que uma aplicação $f : A \to B$ é *injetora* (ou *injetiva*) se elementos distintos de *A* possuem imagens distintas, ou seja,

Se $f(a) = f(a')$, então $a = a'$.

DEFINIÇÃO Dizemos que uma aplicação $f : A \to B$ é *sobrejetora* (ou *sobrejetiva* ou, simplesmente, *sobre*) se cada elemento $b \in B$ é a imagem de, pelo menos, um elemento $a \in A$.

DEFINIÇÃO Uma aplicação $f : A \to B$ é denominada *bijeção* de *A* em *B* (ou *aplicação bijetora* ou *bijetiva*) se *f* for injetora e, também, sobrejetora.

Exemplo 5.3 Sejam $f : \mathbf{R} \to \mathbf{R}$, $g : \mathbf{R} \to \mathbf{R}$, $h : \mathbf{R} \to \mathbf{R}$ definidas por

$$f(x) = 2^x, \qquad g(x) = x^3 - x, \qquad h(x) = x^2$$

Os gráficos dessas funções aparecem na Figura 5-1. A função *f* é injetora. Geometricamente, isso significa que cada reta horizontal não contém mais que um ponto de *f*. A função *g* é sobrejetora. Geometricamente, isso significa que toda reta horizontal contém pelo menos um ponto de *g*. A função *h* não é nem injetora nem sobrejetora. Por exemplo, tanto 2 quanto –2 possuem a mesma imagem 4, e –16 não possui pré-imagem.

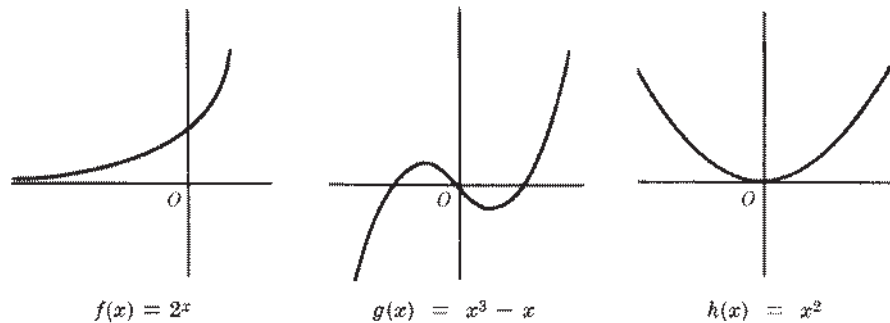



Figura 5-1

Aplicações identidade e inversa

Seja A um conjunto não vazio. A aplicação $f : A \to A$ definida por $f(a) = a$, isto é, a função que associa cada elemento de A a si mesmo, é denominada *aplicação identidade* e, geralmente, é denotada por $\mathbf{1}_A$ ou $\mathbf{1}$ ou I. Assim, para cada $a \in A$, temos $\mathbf{1}_A(a) = a$.

Seja, agora, $f : A \to B$. Dizemos que $g : B \to A$ é a aplicação *inversa* de f, denotada por f^{-1}, se

$$f \circ g = \mathbf{1}_B \qquad \text{e} \qquad g \circ f = \mathbf{1}_A$$

Enfatizamos que f possui uma inversa se, e só se, f é uma bijeção de A em B, isto é, se f é injetora e sobrejetora (Problema 5.7). Além disso, se $b \in B$, então $f^{-1}(b) = a$, onde a é o único elemento de A para o qual $f(a) = b$.

5.3 TRANSFORMAÇÕES LINEARES

Começamos com uma definição.

DEFINIÇÃO Sejam V e U espaços vetoriais sobre um mesmo corpo K. Dizemos que uma aplicação $F : V \to U$ é *linear* ou, mais precisamente, uma *transformação linear*, se satisfizer as duas condições seguintes.

(1) Para quaisquer vetores $v, w \in V$, temos $F(v + w) = F(v) + F(w)$.
(2) Para quaisquer escalar k e vetor $v \in V$, temos $F(kv) = kF(v)$.

Assim, $F : V \to U$ é linear se "preservar" a soma de vetores e a multiplicação por escalar, as duas operações básicas de um espaço vetorial.

Substituindo $k = 0$ na condição (2), obtemos $F(0) = 0$. Assim, qualquer transformação linear leva o vetor nulo no vetor nulo.

Para quaisquer escalares $a, b \in K$ e vetores $v, w \in V$, obtemos

$$F(av + bw) = F(av) + F(bw) = aF(v) + bF(w)$$

De modo mais geral, para quaisquer escalares $a_i \in K$ e vetores $v_i \in V$, obtemos a propriedade básica das transformações lineares, como segue.

$$F(a_1 v_1 + a_2 v_2 + \cdots + a_m v_m) = a_1 F(v_1) + a_2 F(v_2) + \cdots + a_m F(v_m)$$

OBSERVAÇÃO Uma transformação linear $F : V \to U$ fica caracterizada completamente pela condição

$$F(av + bw) = aF(v) + bF(w) \tag{*}$$

portanto, essa condição é, às vezes, utilizada como definição de transformação linear.

Exemplo 5.4

(a) Seja $F : \mathbf{R}^3 \to \mathbf{R}^3$ a aplicação "projeção" sobre o plano xy, isto é, F é a aplicação definida por $F(x, y, z) = (x, y, 0)$. Mostremos que F é linear. Dados $v = (a, b, c)$ e $w = (a', b', c')$, temos

$$\begin{aligned} F(v + w) &= F(a + a',\ b + b',\ c + c') = (a + a',\ b + b',\ 0) \\ &= (a, b, 0) + (a', b', 0) = F(v) + F(w) \end{aligned}$$

e, dado qualquer escalar k,

$$F(kv) = F(ka, kb, kc) = (ka, kb, 0) = k(a, b, 0) = kF(v)$$

Assim, F é linear.

(b) Seja $G : \mathbf{R}^2 \to \mathbf{R}^2$ a aplicação "translação" definida por $G(x, y) = (x + 1, y + 2)$. [Ou seja, G soma o vetor $(1, 2)$ a cada vetor $v = (x, y)$ de $\mathbf{R}^2$.] Observe que

$$G(0) = G(0, 0) = (1, 2) \neq 0$$

Assim, o vetor nulo não é levado no vetor nulo. Portanto, G não é linear.

Exemplo 5.5 (Aplicações derivada e integral) Considere o espaço vetorial $V = \mathbf{P}(t)$ dos polinômios sobre o corpo real **R**. Sejam $u(t)$ e $v(t)$ polinômios quaisquer de V e r um escalar qualquer.

(a) Seja $\mathbf{D}: V \to V$ a aplicação derivada. Demonstra-se no Cálculo que

$$\frac{d(u+v)}{dt} = \frac{du}{dt} + \frac{dv}{dt} \qquad \text{e} \qquad \frac{d(ru)}{dt} = r\frac{du}{dt}$$

Ou seja, $\mathbf{D}(u + v) = \mathbf{D}(u) + \mathbf{D}(v)$ e $\mathbf{D}(ru) = r\mathbf{D}(u)$. Assim, a aplicação derivada é uma transformação linear.

(b) Seja $\mathbf{J}: V \to \mathbf{R}$ uma aplicação integral, digamos,

$$\mathbf{J}(f(t)) = \int_0^1 f(t)\,dt$$

Também se demonstra no Cálculo que

$$\int_0^1 [u(t) + v(t)]dt = \int_0^1 u(t)\,dt + \int_0^1 v(t)\,dt$$

e

$$\int_0^1 ru(t)\,dt = r\int_0^1 u(t)\,dt$$

Ou seja, $\mathbf{J}(u + v) = \mathbf{J}(u) + \mathbf{J}(v)$ e $\mathbf{J}(ru) = r\mathbf{J}(u)$. Assim, a aplicação integral é linear.

Exemplo 5.6 (Aplicações Nula e Identidade)

(a) Seja $F: V \to U$ a aplicação que leva qualquer vetor $v \in V$ no vetor nulo $0 \in U$. Então, dados vetores $v, w \in V$ quaisquer e um escalar $k \in K$ qualquer, temos

$$F(v + w) = 0 = 0 + 0 = F(v) + F(w) \qquad \text{e} \qquad F(kv) = 0 = k0 = kF(v)$$

Assim, F é linear. Dizemos que F é a *transformação nula*, denotada por 0.

(b) Considere a aplicação identidade $I: V \to V$, que leva cada $v \in V$ em si mesmo. Então, dados vetores $v, w \in V$ quaisquer e escalares $a, b \in K$ quaisquer, temos

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Assim, I é linear. Dizemos que I é a *transformação identidade*.

O próximo teorema (demonstrado no Problema 5.13) fornece uma grande quantidade de exemplos de transformações lineares. Em particular, diz que uma transformação linear fica completamente determinada pelos seus valores nos elementos de alguma base.

Teorema 5.2 Sejam V e U espaços vetoriais sobre um corpo K. Sejam $\{v_1, v_2, ..., v_n\}$ uma base de V e $u_1, u_2, ..., u_n$ vetores quaisquer de U. Então existe uma única transformação linear $F: V \to U$ tal que $F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

Enfatizamos que os vetores $u_1, u_2, ..., u_n$ nesse teorema são completamente arbitrários, podendo ser linearmente dependentes ou, até mesmo, todos iguais entre si.

Matrizes como transformações lineares

Seja A uma matriz real $m \times n$. Vimos que A determina uma aplicação matricial $F_A: K^n \to K^m$ dada por $F_A(u) = Au$ (em que os vetores de K^n e K^m são escritos como colunas). Mostremos que F_A é linear. Usando a multiplicação matricial, temos

$$F_A(v + w) = A(v + w) = Av + Aw = F_A(v) + F_A(w)$$
$$F_A(kv) = A(kv) = k(Av) = kF_A(v)$$

Em outras palavras, usando A para representar a aplicação matricial, temos

$$A(v + w) = Av + Aw \qquad \text{e} \qquad A(kv) = k(Av)$$

Assim, A é linear e costumamos dizer que A é uma transformação matricial.

Isomorfismo de espaços vetoriais

A noção de isomorfismo entre dois espaços vetoriais já foi definida no Capítulo 4, quando estudamos as coordenadas de um vetor em relação a uma base. Agora revemos esse conceito.

DEFINIÇÃO Dois espaços vetoriais V e U sobre K são *isomorfos*, e escrevemos $V \cong U$, se existir uma transformação linear bijetora $F : V \to U$. Nesse caso, dizemos que F é um *isomorfismo* entre V e U.

Considere um espaço vetorial V de dimensão n qualquer e seja S uma base qualquer de V. Então a aplicação

$$v \mapsto [v]_S$$

que leva cada vetor $v \in V$ em seu vetor de coordenadas $[v]_S$ é um isomorfismo entre V e K^n.

5.4 NÚCLEO E IMAGEM DE UMA TRANSFORMAÇÃO LINEAR

Começamos definindo dois conceitos.

DEFINIÇÃO Seja $F : V \to U$ uma transformação linear. O *núcleo* de F, denotado por Nuc F, é o conjunto de todos os elementos de V que são levados no vetor nulo 0 de U, ou seja,

$$\text{Nuc } F = \{v \in V : F(v) = 0\}$$

A *imagem* de F, denotada por Im F, é o conjunto dos pontos imagens de U, ou seja,

$$\text{Im } F = \{u \in U : \text{existe } v \in V \text{ tal que } F(v) = u\}$$

O teorema abaixo é facilmente demonstrado (Problema 5.22).

Teorema 5.3 Seja $F : V \to U$ uma transformação linear. Então o núcleo de F é um subespaço de V e a imagem de F é um subespaço de U.

Agora suponha que $v_1, v_2, ..., v_m$ gerem um espaço vetorial V e que $F : V \to U$ seja uma transformação linear. Mostremos que $F(v_1), F(v_2), ..., F(v_m)$ geram Im F. Seja $u \in \text{Im } F$. Então existe algum $v \in V$ tal que $F(v) = u$. Como os v_i geram V e como $v \in V$, existem escalares $a_1, a_2, ..., a_m$ para os quais

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m$$

Portanto,

$$u = F(v) = F(a_1 v_1 + a_2 v_2 + \cdots + a_m v_m) = a_1 F(v_1) + a_2 F(v_2) + \cdots + a_m F(v_m)$$

Assim, os vetores $F(v_1), F(v_2), ..., F(v_m)$ geram Im F.

Enunciamos esse resultado formalmente.

Proposição 5.4 Se $v_1, v_2, ..., v_m$ geram um espaço vetorial V e $F : V \to U$ é uma transformação linear, então $F(v_1), F(v_2), ..., F(v_m)$ geram Im F.

Exemplo 5.7

(a) Seja $F : \mathbf{R}^3 \to \mathbf{R}^3$ a projeção de um vetor v sobre o plano xy [conforme Figura 5-2(a)], ou seja,

$$F(x, y, z) = (x, y, 0)$$

Claramente, a imagem de F é todo o plano xy, ou seja, os pontos da forma $(x, y, 0)$. Além disso, o núcleo de F é o eixo z, ou seja,

$$\text{Im } F = \{(a, b, c) : c = 0\} = \text{plano } xy \quad \text{e} \quad \text{Nuc } F = \{(a, b, c) : a = 0, b = 0\} = \text{eixo } z$$

(b) Seja $G : \mathbf{R}^3 \to \mathbf{R}^3$ a transformação linear que gira cada vetor v em torno do eixo z por um ângulo θ [conforme Figura 5-2(b)], ou seja,

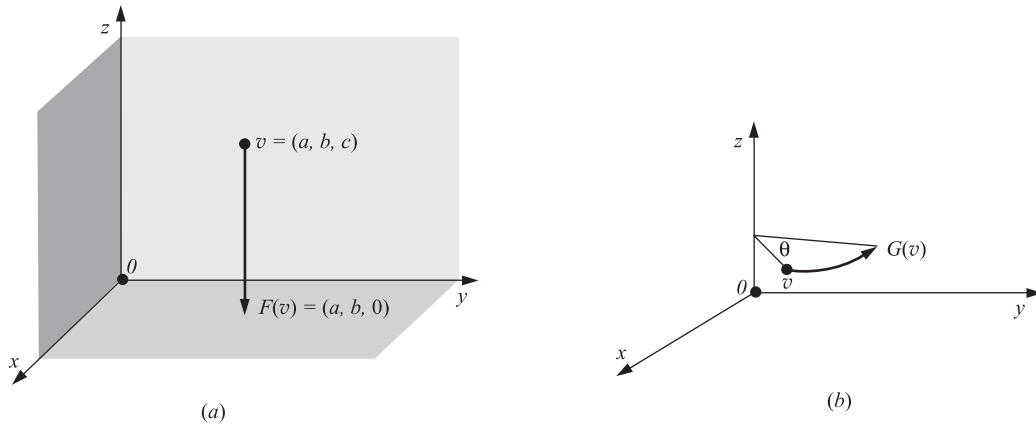



Figura 5-2

$$G(x, y, z) = (x \cos\theta - y \operatorname{sen}\theta,\ x \operatorname{sen}\theta + y \cos\theta,\ z)$$

Observe que a distância de um vetor v até a origem O não muda com essa rotação, portanto, apenas o vetor nulo 0 é levado no vetor nulo 0. Assim, Nuc $G = \{0\}$. Por outro lado, cada vetor $\mathbf{R}^3$ é a imagem de um vetor v de $\mathbf{R}^3$, que pode ser obtido girando u de volta pelo mesmo ângulo θ. Assim, Im $G = \mathbf{R}^3$, todo o espaço.

Exemplo 5.8 Considere o espaço vetorial $V = \mathbf{P}(t)$ dos polinômios sobre o corpo real $\mathbf{R}$ e seja $H : V \to V$ o operador derivada terceira, isto é, $H[f(t)] = d^3f/dt^3$. [Às vezes se utiliza a notação $\mathbf{D}^3$ para H, onde $\mathbf{D}$ é o operador derivada.] Afirmamos que

$$\text{Nuc } H = \{\text{polinômios de grau} \leq 2\} = \mathbf{P}_2(t) \qquad \text{e} \qquad \text{Im } H = V$$

A primeira afirmação decorre de $H(at^2 + bt + c) = 0$, mas $H(t^n) \neq 0$, para $n \geq 3$. A segunda decorre do fato de que qualquer polinômio $g(t)$ de V é a derivada terceira de algum polinômio $f(t)$ (que pode ser encontrado tomando a antiderivada de $g(t)$ três vezes).

Núcleo e imagem de transformações matriciais

Sejam A uma matriz 3×4 e $\{e_1, e_2, e_3, e_4\}$ a base canônica de K^4 (escrita como colunas).

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix}, \qquad e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad e_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad e_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Vimos que A pode ser interpretada como uma transformação linear $A : K^4 \to K^3$, em que os vetores de K^4 e K^3 são considerados vetores colunas. Como a base canônica gera K^4, as imagens Ae_1, Ae_2, Ae_3, Ae_4 geram a imagem de A. Mas os vetores Ae_1, Ae_2, Ae_3, Ae_4 são exatamente as colunas de A, como segue.

$$Ae_1 = [a_1, b_1, c_1]^T, \qquad Ae_2 = [a_2, b_2, c_2]^T, \qquad Ae_3 = [a_3, b_3, c_3]^T, \qquad Ae_4 = [a_4, b_4, c_4]^T$$

Assim, a imagem de A é exatamente o espaço coluna de A.

Por outro lado, o núcleo de A consiste em todos os vetores v para os quais $Av = 0$. Isso significa que o núcleo de A é o espaço solução do sistema homogêneo $AX = 0$, denominado *espaço nulo* de A.

Enunciamos esses resultados formalmente.

Proposição 5.5 Seja A uma matriz $m \times n$ sobre o corpo K interpretada como uma transformação linear $A : K^n \to K^m$. Então

$$\text{Nuc } A = \text{nul}(A) \qquad \text{e} \qquad \text{Im } A = \text{col}(A)$$

Aqui, col(A) denota o espaço coluna de A e nul(A) o espaço nulo de A.

Posto e nulidade de uma transformação linear

Seja $F : V \to U$ uma transformação linear. O *posto* de F é definido como a dimensão da imagem de F e a *nulidade* de F é definida como a dimensão do núcleo de F, ou seja,

$$\text{pos}(F) = \dim(\text{Im } F) \qquad \text{e} \qquad \text{nul}(F) = \dim(\text{Nuc } F)$$

Vale o teorema seguinte (demonstrado no Problema 5.23).

Teorema 5.6 Sejam V um espaço de dimensão finita e $F : V \to U$ uma transformação linear. Então

$$\dim V = \dim(\text{Nuc } F) + \dim(\text{Im } F) = \text{nul}(F) + \text{pos}(F)$$

Lembramos que o posto de uma matriz A também foi definido como a dimensão dos espaços coluna e linha de A. Se, agora, interpretamos A como uma transformação linear, então ambas definições coincidem, pois a imagem de A é, precisamente, o espaço coluna de A.

Exemplo 5.9 Seja $F : \mathbf{R}^4 \to \mathbf{R}^3$ a transformação linear definida por

$$F(x, y, z, t) = (x - y + z + t, \quad 2x - 2y + 3z + 4t, \quad 3x - 3y + 4z + 5t)$$

(a) Encontre uma base e a dimensão da imagem de F.

Primeiro calculamos a imagem dos vetores da base canônica de $\mathbf{R}^4$.

$$F(1,0,0,0) = (1,2,3), \qquad F(0,0,1,0) = (1,3,4)$$
$$F(0,1,0,0) = (-1,-2,-3), \qquad F(0,0,0,1) = (1,4,5)$$

Pela Proposição 5.4, os vetores imagem geram Im F. Portanto, formamos a matriz M cujas linhas são esses vetores imagem e reduzimos M à forma escalonada, como segue.

$$M = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 1 & 3 & 4 \\ 1 & 4 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, (1, 2, 3) e (0, 1, 1) formam uma base de Im F. Portanto, $\dim(\text{Im } F) = 2$ e $\text{pos}(F) = 2$.

(b) Encontre uma base e a dimensão do núcleo da transformação F.

Tomamos $F(v) = 0$, com $v = (x, y, z, t)$, obtendo

$$F(x, y, z, t) = (x - y + z + t, \quad 2x - 2y + 3z + 4t, \quad 3x - 3y + 4z + 5t) = (0, 0, 0)$$

Igualando os componentes correspondentes, formamos o sistema homogêneo seguinte, cujo espaço solução é o núcleo de F.

$$\begin{aligned} x - y + z + t &= 0 \\ 2x - 2y + 3z + 4t &= 0 \\ 3x - 3y + 4z + 5t &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x - y + z + t &= 0 \\ z + 2t &= 0 \\ z + 2t &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x - y + z + t &= 0 \\ z + 2t &= 0 \end{aligned}$$

As variáveis livres são y e t. Portanto, $\dim(\text{Nuc } F) = 2$, ou $\text{nul}(F) = 2$.

(i) Tomando $y = 1$, $t = 0$ obtemos a solução (–1, 1, 0, 0).
(ii) Tomando $y = 0$, $t = 1$ obtemos a solução (1, 0, –2, 1).

Assim (–1, 1, 0, 0) e (1, 0, –2, 1) formam uma base de Nuc F.

Conforme afirma o Teorema 5.6, temos $\dim(\text{Im F}) + \dim(\text{Nuc F}) = 4 = \dim \mathbf{R}^4$.

Aplicações a sistemas de equações lineares

Seja $AX = B$ a forma matricial de um sistema de m equações lineares com n incógnitas. A matriz A pode ser considerada como uma transformação linear

$$A : K^n \to K^m$$

Assim, a solução da equação $AX = B$ pode ser interpretada como a pré-imagem do vetor $B \in K^m$ pela transformação linear A. Além disso, o espaço solução do sistema homogêneo associado

$$AX = 0$$

pode ser interpretado como o núcleo da transformação linear A. Aplicando o Teorema 5.6 a esse sistema homogêneo, obtemos

$$\dim(\text{Nuc } A) = \dim K^n - \dim(\text{Im } A) = n - \text{pos}(A)$$

Mas n é exatamente o número de incógnitas do sistema homogêneo $AX = 0$. Assim, demonstramos o seguinte teorema do Capítulo 4.

Teorema 4.19 A dimensão do espaço solução W de um sistema homogêneo $AX = 0$ é $n - r$, onde n é o número de incógnitas e r é o posto da matriz de coeficientes A.

Observe que r também é o número de variáveis pivô na forma escalonada de $AX = 0$, portanto, $s = n - r$ também é o número de variáveis livres. Além disso, são linearmente independentes os s vetores solução de $AX = 0$ dados no Teorema 3.14 (Problema 4.52). Dessa forma, como $\dim W = s$, esses vetores formam uma base do espaço solução W. Assim, demonstramos também o Teorema 3.14.

5.5 TRANSFORMAÇÕES LINEARES SINGULARES E NÃO SINGULARES, ISOMORFISMOS

Seja $F: V \to U$ uma transformação linear. Vimos que, sempre, $F(0) = 0$. Dizemos que F é *singular* se a imagem de algum vetor não nulo v for 0, isto é, se existir $v \neq 0$ tal que $F(v) = 0$. Assim, $F: V \to U$ é *não singular* se o vetor nulo 0 for o único vetor cuja imagem por F é 0; em outras palavras, se $\text{Nuc } F = \{0\}$.

Exemplo 5.10 Considere a projeção $F: \mathbf{R}^3 \to \mathbf{R}^3$ e a aplicação rotação $G: \mathbf{R}^3 \to \mathbf{R}^3$ representadas na Figura 5-2. (Ver Exemplo 5.7.) Como o núcleo de F é o eixo z, F é singular. Por outro lado, o núcleo de G é composto apenas pelo vetor nulo. Assim, G é não singular.

Transformações não singulares podem ser também caracterizadas como as transformações que levam conjuntos independentes em conjuntos independentes. Mais precisamente, demonstra-se (Problema 5.28) o teorema seguinte.

Teorema 5.7 Seja $F: V \to U$ uma transformação linear não singular. Então a imagem de qualquer conjunto linearmente independente é linearmente independente.

Isomorfismos

Suponha que a transformação linear $F: V \to U$ seja injetora. Então apenas $0 \in V$ pode ser levado em $0 \in U$ e, portanto, F é não singular. A recíproca também é verdadeira. De fato, supondo que F seja não singular e que $F(v) = F(w)$, então $F(v - w) = F(v) - F(w) = 0$ e, portanto, $v - w = 0$ e $v = w$. Assim, $F(v) = F(w)$ implica $v = w$, ou seja, F é injetora. Dessa forma, demonstramos a proposição seguinte.

Proposição 5.8 Uma transformação linear $F: V \to U$ é injetora se, e só se, F é não singular.

Vimos que uma aplicação $F: V \to U$ é dita um *isomorfismo* se F for linear e bijetora, isto é, se F for injetora e sobrejetora. Também dizemos que um espaço vetorial V é *isomorfo* a um espaço vetorial U, e escrevemos $V \cong U$, se existir um isomorfismo $F: V \to U$.

Vale o seguinte teorema (demonstrado no Problema 5.29).

Teorema 5.9 Seja V um espaço vetorial de dimensão finita e suponha que $\dim V = \dim U$. Seja $F: V \to U$ linear. Então F é um isomorfismo se, e só se, F é não singular.

5.6 OPERAÇÕES COM TRANSFORMAÇÕES LINEARES

Podemos combinar transformações lineares de várias maneiras para obter novas transformações lineares. Essas operações são muito importantes e serão utilizadas em todo este texto.

Sejam $F: V \to U$ e $G: V \to U$ transformações lineares sobre um corpo K. A soma $F + G$ e a multiplicação por escalar kF, com $k \in K$, são definidas como as transformações lineares de V em U dadas por

$$(F+G)(v) \equiv F(v) + G(v) \qquad \text{e} \qquad (kF)(v) \equiv kF(v)$$

Mostremos, agora, que, se F e G são lineares, então $F + G$ e kF também são lineares. De fato, dados vetores $v, w \in V$ e escalares $a, b \in K$ quaisquer, temos

$$\begin{aligned}(F+G)(av+bw) &= F(av+bw) + G(av+bw)\\ &= aF(v) + bF(w) + aG(v) + bG(w)\\ &= a[F(v)+G(v)] + b[F(w)+G(w)]\\ &= a(F+G)(v) + b(F+G)(w)\end{aligned}$$

e

$$\begin{aligned}(kF)(av+bw) &= kF(av+bw) = k[aF(v)+bF(w)]\\ &= akF(v) + bkF(w) = a(kF)(v) + b(kF)(w)\end{aligned}$$

Assim, $F + G$ e kF são lineares.

Vale o teorema a seguir.

Teorema 5.10 Sejam V e U espaços vetoriais sobre um corpo K. Então a coleção de todas as transformações lineares de V em U com as operações de soma e multiplicação por escalar constitui um espaço vetorial sobre K.

O espaço vetorial das transformações lineares do Teorema 5.10 é geralmente denotado por

$$\text{Hom}(V, U)$$

Aqui, "hom" vem da palavra "homomorfismo". Enfatizamos que a demonstração do Teorema 5.10 reduz-se a demonstrar que Hom(V, U) efetivamente satisfaz os oito axiomas de um espaço vetorial. O elemento nulo de Hom(V, U) é a *transformação nula* de V em U, denotada por 0 e definida por

$$0(v) = 0$$

para cada vetor $v \in V$.

Para espaços vetoriais V e U de dimensão finita temos o teorema seguinte.

Teorema 5.11 Se dim $V = m$ e dim $U = n$, então dim[Hom(V, U)] $= mn$.

Composição de transformações lineares

Sejam, agora, V, U e W espaços vetoriais sobre um mesmo corpo K e sejam $F: V \to U$ e $G: U \to W$ transformações lineares. Podemos visualizar essas transformações como segue.

$$V \xrightarrow{F} U \xrightarrow{G} W$$

Já definimos a função composta $G \circ F$ como a aplicação de V em W definida por $(G \circ F)(v) = G(F(v))$. Mostremos que $G \circ F$ é linear sempre que F e G forem lineares. De fato, dados vetores $v, w \in V$ e escalares $a, b \in K$ quaisquer, temos

$$\begin{aligned}(G\circ F)(av+bw) &= G(F(av+bw)) = G(aF(v)+bF(w))\\ &= aG(F(v)) + bG(F(w)) = a(G\circ F)(v) + b(G\circ F)(w)\end{aligned}$$

Assim, $G \circ F$ é linear.

A composição de transformações lineares e as operações de soma e multiplicação por escalar estão relacionadas como segue.

Teorema 5.12 Se V, U e W são espaços vetoriais sobre K e

$$F: V \to U, \qquad F': V \to U \qquad \text{e} \qquad G: U \to W, \qquad G': U \to W$$

são transformações lineares, então

(i) $G \circ (F + F') = G \circ F + G \circ F'$,
(ii) $(G + G') \circ F = G \circ F + G' \circ F$,
(iii) $k(G \circ F) = (kG) \circ F = G \circ (kF)$, com qualquer escalar $k \in K$.

5.7 A ÁLGEBRA *A*(*V*) DOS OPERADORES LINEARES

Seja V um espaço vetorial sobre um corpo K. Nesta seção consideramos o caso particular das transformações lineares do espaço vetorial V nele mesmo, ou seja, das transformações lineares da forma $F: V \to V$. Essas transformações lineares de V em V são denominadas *operadores lineares* de V. Escrevemos $A(V)$, em vez de Hom(V, V), para o espaço de todos os operadores de V.

Em particular, $A(V)$ é um espaço vetorial sobre K (Teorema 5.10) e, se dim $V = n$, então dim $A(V) = n^2$. Além disso, dados operadores $F, G \in A(V)$, existe a composta $G \circ F$, que também pertence a $A(V)$. Dessa forma, temos uma "multiplicação" em $A(V)$. [Às vezes, escrevemos GF em vez de $G \circ F$ no espaço $A(V)$.]

OBSERVAÇÃO Uma *álgebra* A sobre um corpo K é um espaço vetorial sobre K no qual está definida uma operação de multiplicação que satisfaz, para $F, G, H \in A$ e $k \in K$ quaisquer,

(i) $F(G + H) = FG + FH$,
(ii) $(G + H)F = GF + HF$,
(iii) $k(GF) = (kG)F = G(kF)$

Dizemos que uma álgebra é *associativa* se, além disso, $(FG)H = F(GH)$.

A definição de álgebra e teoremas precedentes fornecem o resultado seguinte.

Teorema 5.13 Seja V um espaço vetorial sobre K. Então $A(V)$ é uma álgebra associativa sobre K em relação à composição de operadores. Se dim $V = n$, então dim $A(V) = n^2$.

Por essa razão dizemos que $A(V)$ é a *álgebra dos operadores lineares* de V.

Polinômios e operadores lineares

Observe que a transformação identidade $I: V \to V$ pertence a $A(V)$. Também, dado qualquer operador linear F em $A(V)$, temos $FI = IF = F$. Podemos, também, definir "potências" de F, como segue.

$$F^0 = I, \qquad F^2 = F \circ F, \qquad F^3 = F^2 \circ F = F \circ F \circ F, \qquad F^4 = F^3 \circ F, \qquad \ldots$$

Além disso, dado qualquer polinômio $p(t)$ sobre K, digamos,

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^2$$

podemos formar o operador linear $p(F)$, definido por

$$p(F) = a_0 I + a_1 F + a_2 F^2 + \cdots + a_s F^s$$

(Dado um escalar k qualquer, o operador kI é, às vezes, denotado simplesmente por k.) Em particular, dizemos que F é um *zero* do polinômio $p(t)$ se $p(F) = 0$.

Exemplo 5.11 Seja $F: K^3 \to K^3$ definido por $F(x, y, z) = (0, x, y)$. Dado qualquer $(a, b, c) \in K^3$, temos

$$(F + I)(a, b, c) = (0, a, b) + (a, b, c) = (a,\ a + b,\ b + c)$$
$$F^3(a, b, c) = F^2(0, a, b) = F(0, 0, a) = (0, 0, 0)$$

Assim, $F^3 = 0$, o operador nulo de $A(V)$. Isso significa que F é um zero do polinômio $p(t) = t^3$.

Matrizes quadradas como operadores lineares

Seja $\mathbf{M} = \mathbf{M}_{n,n}$ o espaço vetorial de todas as matrizes quadradas de ordem n sobre K. Então qualquer matriz A em $\mathbf{M}$ define uma transformação linear $F_A : K^n \to K^n$ por meio de $F_A(u) = Au$ (escrevendo os vetores de K^n como colunas). Como essa transformação é de K^n em si mesmo, a matriz quadrada é um operador linear e não apenas uma transformação linear.

Se A e B são matrizes de $\mathbf{M}$, então o produto AB está definido. Além disso, dado qualquer vetor (coluna) u de K^n,

$$F_{AB}(u) = (AB)u = A(Bu) = A(F_B(U)) = F_A(F_B(u)) = (F_A \circ F_B)(u)$$

Em outras palavras, o produto matricial AB corresponde à composição de A e B como transformações lineares. Analogamente, a soma matricial $A + B$ corresponde à soma de A e B como transformações lineares e a multiplicação por escalar kA corresponde ao múltiplo escalar de kA de A como uma transformação linear.

Operadores invertíveis em $A(V)$

Seja $F : V \to V$ um operador linear. Dizemos que F é *invertível* se possuir uma inversa, ou seja, se existir F^{-1} em $A(V)$ tal que $FF^{-1} = F^{-1}F = I$. Por outro lado, F é invertível como uma aplicação se for injetora e sobre. Nesse caso, a aplicação inversa F^{-1} também é linear e é a inversa de F como operador linear (demonstrado no Problema 5.15).

Suponha que F seja invertível. Então apenas $0 \in V$ pode ser levado em si mesmo e, portanto, F é não singular. A recíproca não é verdadeira, conforme exemplo a seguir.

Exemplo 5.12 Seja $V = \mathbf{P}(t)$ o espaço vetorial dos polinômios sobre K. Seja F o operador de V que aumenta uma unidade em cada expoente de t, ou seja,

$$F(a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s) = a_0 t + a_1 t^2 + a_2 t^3 + \cdots + a_s t^{s+1}$$

Então, F é um operador não singular de V. Contudo, F não é sobre, portanto, F não é invertível.

O espaço vetorial $V = \mathbf{P}(t)$ desse exemplo é de dimensão infinita. A situação muda completamente se V tem dimensão finita, conforme teorema a seguir.

Teorema 5.14 Seja F um operador linear de um espaço vetorial V de dimensão finita. Então as quatro condições seguintes são equivalentes.

(i) F é não singular: Nuc $F = \{0\}$.
(ii) F é uma aplicação injetora.
(iii) F é uma aplicação sobre.
(iv) F é um operador invertível.

A demonstração desse teorema decorre, essencialmente, do Teorema 5.6, que afirma que

$$\dim V = \dim(\text{Nuc } F) + \dim(\text{Im } F)$$

Pela Proposição 5.8, (i) e (ii) são equivalentes. Observe que (iv) é equivalente a (ii) e (iii). Assim, para demonstrar o teorema, basta provar que (i) e (iii) são equivalentes, conforme segue.

(a) Suponha que (i) seja válida. Então $\dim(\text{Nuc } F) = 0$, portanto, a equação fornece $\dim V = \dim(\text{Im F})$, o que significa que $V = \text{Im } F$, ou seja, que F é sobrejetora. Assim, (i) implica (iii).
(b) Suponha que (iii) seja válida. Então $V = \text{Im } F$ e, portanto, $\dim V = \dim(\text{Im } F)$. Agora a equação fornece $\dim(\text{Nuc } F) = 0$ e, portanto, F é não singular. Assim, (iii) implica (i).

Dessa forma, as quatro condições do teorema são equivalentes.

OBSERVAÇÃO Suponha que A seja uma matriz quadrada de ordem n sobre K. Então A pode ser vista como um operador linear de K^n. Como K^n tem dimensão finita, vale o Teorema 5.14 para a matriz quadrada A. Esse é o motivo pelo qual podemos usar os termos "não singular" e "invertível" sem distinção quando aplicados a matrizes quadradas.

Exemplo 5.13 Seja F o operador linear de $\mathbf{R}^2$ definido por $F(x, y) = (2x + y, 3x + 2y)$.

(a) Para mostrar que F é invertível, basta mostrar que F é não singular. Tomamos $F(x, y) = (0, 0)$ e obtemos o sistema homogêneo

$$2x + y = 0 \quad \text{e} \quad 3x + 2y = 0$$

Resolvendo em x e y, obtemos $x = 0$ e $y = 0$. Logo, F é não singular e, portanto, invertível.

(b) Para determinar uma fórmula para F^{-1}, tomamos $F(x, y) = (s, t)$ e, portanto, $F^{-1}(s, t) = (x, y)$. Temos

$$(2x + y,\ \ 3x + 2y) = (s, t) \qquad \text{ou} \qquad \begin{aligned} 2x + \ \ y &= s \\ 3x + 2y &= t \end{aligned}$$

Resolvendo x e y em termos de s, t, obtemos $x = 2s - t$ e $y = -3s + 2t$. Assim,

$$F^{-1}(s, t) = (2s - t,\ \ -3s + 2t) \qquad \text{ou} \qquad F^{-1}(x, y) = (2x - y,\ \ -3x + 2y)$$

em que reescrevemos a fórmula para F^{-1} usando x e y em vez de s e t.

Problemas Resolvidos

Aplicações

5.1 Decida se os diagramas da Figura 5-3 definem aplicações de $A = \{a, b, c\}$ em $B = \{x, y, z\}$.

(a) Não. Não há coisa alguma associada ao elemento $b \in A$.

(b) Não. Os dois elementos x e z estão associados a $c \in A$.

(c) Sim.

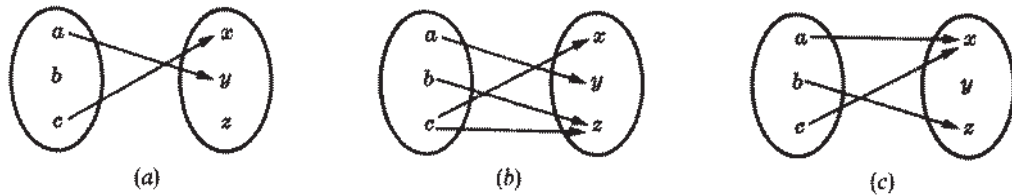



Figura 5-3

5.2 Sejam $f : A \to B$ e $g : B \to C$ definidas pela Figura 5-4.

(a) Encontre a aplicação composta $(g \circ f) : A \to C$.

(b) Encontre as imagens das aplicações f, g, $g \circ f$.

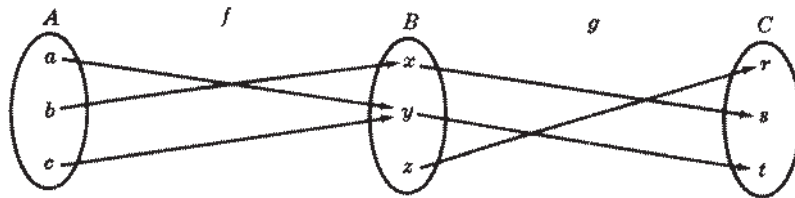



Figura 5-4

(a) Usando a definição de aplicação composta, calculamos

$$(g \circ f)\ (a) = g(f(a)) = g(y) = t, \qquad (g \circ f)\ (b) = g(f(b)) = g(x) = s$$
$$(g \circ f)\ (c) = g(f(c)) = g(y) = t$$

Observe que alcançamos o mesmo resultado "seguindo as setas" na Figura 5-4.

$$a \to y \to t, \qquad b \to x \to s, \qquad c \to y \to t$$

(b) Pela Figura 5-4, os valores da imagem de f são x e y e os valores da imagem de g são r, s, t. Logo,

$$\text{Im } f = \{x, y\} \qquad \text{e} \qquad \text{Im } g = \{r, s, t\}$$

Também, pela parte (a), os valores da imagem da composta $g \circ f$ são t e s. Logo, Im $g \circ f = \{s, t\}$. Observe que as imagens de g e $g \circ f$ são diferentes.

5.3 Considere a aplicação $F : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (yz, x^2)$. Encontre

(a) $F(2, 3, 4)$; (b) $F(5, -2, 7)$; (c) $F^{-1}(0, 0)$, ou seja, todos $v \in \mathbf{R}^3$ tais que $F(v) = 0$.

(a) Substituímos na fórmula de F para obter $F(2, 3, 4) = (3 \cdot 4, 2^2) = (12, 4)$.

(b) $F(5, -2, 7) = (-2 \cdot 7, 5^2) = (-14, 25)$.

(c) Tomemos $F(v) = 0$, com $v = (x, y, z)$ e resolvamos em x, y, z.

$$F(x, y, z) = (yz, x^2) = (0, 0) \qquad \text{ou} \qquad yz = 0, x^2 = 0$$

Assim, $x = 0$ e $y = 0$, ou $z = 0$. Em outras palavras, $x = 0$, $y = 0$ ou $x = 0$, $z = 0$, isto é, o eixo z unido com o eixo y.

5.4 Considere a aplicação $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (3y, 2x)$. Seja S o círculo unitário de $\mathbf{R}^2$, ou seja, o conjunto solução de $x^2 + y^2 = 1$. (a) Descreva $F(S)$; (b) Encontre $F^{-1}(S)$.

(a) Seja (a, b) um elemento de $F(S)$. Então existe algum $(x, y) \in S$ tal que $F(x, y) = (a, b)$. Logo,

$$(3y, 2x) = (a, b) \qquad \text{ou} \qquad 3y = a, 2x = b \qquad \text{ou} \qquad y = \frac{a}{3}, x = \frac{b}{2}$$

Como $(x, y) \in S$, ou seja, $x^2 + y^2 = 1$, temos

$$\left(\frac{b}{2}\right)^2 + \left(\frac{a}{3}\right)^2 = 1 \qquad \text{ou} \qquad \frac{a^2}{9} + \frac{b^2}{4} = 1$$

Assim, $F(S)$ é uma elipse.

(b) Seja $F(x, y) = (a, b)$, com $(a,b) \in S$. Então $(3y, 2x) = (a, b)$, ou $3y = a$, $2x = b$. Como $(a, b) \in S$, temos $a^2 + b^2 = 1$ Assim, $(3y)^2 + (2x)^2 = 1$ e, portanto, $F^{-1}(S)$ é a elipse $4x^2 + 9y^2 = 1$.

5.5 Sejam as aplicações $f : A \to B$, $g : B \to C$, $h : C \to D$ definidas pela Figura 5-5. Decida se cada uma dessas aplicações é (a) injetora; (b) sobre; (c) invertível (ou seja, possui inversa).

(a) A aplicação $f : A \to B$ é injetora, pois cada elemento tem uma imagem diferente. A aplicação $g : B \to C$ não é injetora porque ambos x e z têm a mesma imagem 4. A aplicação $h : C \to D$ é injetora.

(b) A aplicação $f : A \to B$ não é sobre, porque $z \in B$ não é a imagem de elemento algum de A. A aplicação $g : B \to C$ é sobre, pois cada elemento de C é a imagem de algum elemento de B. A aplicação $h : C \to D$ também é sobrejetora.

(c) Uma aplicação tem uma inversa se, e só se, é injetora e sobre. Logo, somente h tem uma inversa.

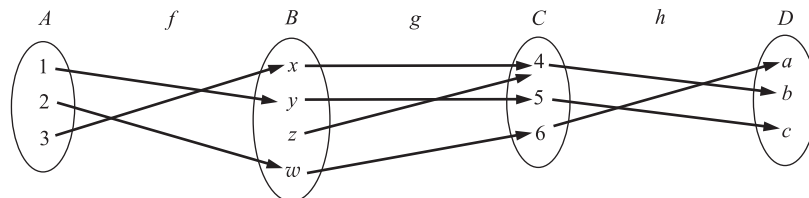



Figura 5-5

5.6 Sejam $f: A \to B$ e $g: B \to C$, portanto, existe $(g \circ f): A \to C$. Demonstre.

(a) Se f e g são injetoras, então $g \circ f$ é injetora.

(b) Se f e g são sobrejetoras, então $g \circ f$ é sobrejetora.

(c) Se $g \circ f$ é injetora, então f é injetora.

(d) Se $g \circ f$ é sobrejetora, então g é sobrejetora.

(a) Suponha que $(g \circ f)(x) = (g \circ f)(y)$. Então $g(f(x)) = g(f(y))$. Como g é injetora, $f(x) = f(y)$. Como f é injetora, $x = y$. Assim, provamos que $(g \circ f)(x) = (g \circ f)(y)$ implica $x = y$ e, portanto, $x = y$ é injetora.

(b) Seja $c \in C$. Como g é sobre, existe $b \in B$ para o qual $g(b) = c$. Como f é sobre, existe $a \in A$ para o qual $f(a) = b$. Assim, $(g \circ f)(a) = g(f(a)) = g(b) = c$ e, portanto, $g \circ f$ é sobre.

(c) Suponha que f não seja injetora. Então existem elementos distintos $x, y \in A$ para os quais $f(x) = f(y)$. Assim, $(g \circ f)(x) = g(f(x)) = g(f(y)) = (g \circ f)(y)$ e, portanto, $g \circ f$ não é injetora. Isso prova que, se $g \circ f$ é injetora, então f necessariamente é injetora.

(d) Dado $a \in A$, temos $(g \circ f)(a) = g(f(a)) \in g(B)$. Logo, $(g \circ f)(A) \subseteq g(B)$. Supondo que g não seja sobrejetora, $g(B)$ estará contido propriamente em C e, portanto, $(g \circ f)(A)$ estará contido propriamente em C, de modo que $g \circ f$ não será sobre. Isso prova que, se $g \circ f$ é sobre, então g é sobre.

5.7 Prove que $f: A \to B$ possui uma inversa se, e só se, f é injetora e sobre.

Suponha que f possua uma inversa, ou seja, que exista uma função $f^{-1}: B \to A$ para a qual $f^{-1} \circ f = \mathbf{1}_A$ e $f \circ f^{-1} = \mathbf{1}_B$. Como $\mathbf{1}_A$ é injetora, f é injetora pelo Problema 5.6(c) e, como $\mathbf{1}_B$ é sobre, f é sobre pelo Problema 5.6(d). Assim, f é injetora e sobrejetora.

Agora suponha que f seja injetora e sobrejetora. Então cada $b \in B$ é a imagem de um único elemento de A, digamos, b^*. Assim, se $f(a) = b$ então $a = b^*$; logo, $f(b^*) = b$. Seja, agora, g a aplicação de B para A definida por $b \mapsto b^*$. Temos

(i) $(g \circ f)(a) = g(f(a)) = g(b) = b^* = a$ para cada $a \in A$ e, portanto, $g \circ f = \mathbf{1}_A$.

(ii) $(f \circ g)(b) = f(g(b)) = f(b^*) = b$ para cada $b \in B$ e, portanto, $f \circ g = \mathbf{1}_B$.

Isso mostra que f possui uma inversa, a saber, a aplicação g.

5.8 Seja $f: \mathbf{R} \to \mathbf{R}$ definida por $f(x) = 2x - 3$. Ocorre que f é injetora e sobre, portanto, possui uma aplicação inversa f^{-1}. Encontre uma fórmula para f^{-1}.

Seja y a imagem de x pela aplicação f, ou seja, $y = 2x - 3$. Então x é a imagem de y pela aplicação inversa f^{-1}. Assim, resolvemos essa equação para x em termos de y para obter $x = \frac{1}{2}(y + 3)$. Então a fórmula que define a função inversa é $f^{-1}(y) = \frac{1}{2}(y + 3)$ ou, usando x em vez de y, $f^{-1}(x) = \frac{1}{2}(x + 3)$.

Transformações lineares

5.9 Seja $F: \mathbf{R}^2 \to \mathbf{R}^2$ a aplicação definida por $F(x, y) = (x + y, x)$. Mostre que F é linear.

Devemos mostrar que $F(v + w) = F(v) + F(w)$ e que $F(kv) = kF(v)$, onde v e w são elementos quaisquer de $\mathbf{R}^2$ e k é um escalar qualquer. Sejam $v = (a, b)$ e $w = (a', b')$. Então

$$v + w = (a + a',\ b + b') \qquad \text{e} \qquad kv = (ka, kb)$$

Temos $F(v) = (a + b, a)$ e $F(w) = (a' + b', a')$. Assim,

$$\begin{aligned} F(v + w) &= F(a + a',\ b + b') = (a + a' + b + b',\ a + a') \\ &= (a + b,\ a) + (a' + b',\ a') = F(v) + F(w) \end{aligned}$$

e

$$F(kv) = F(ka, kb) = (ka + kb,\ ka) = k(a + b,\ a) = kF(v)$$

Como v, w, k são arbitrários, temos que F é linear.

5.10 Seja $F : \mathbf{R}^3 \to \mathbf{R}^2$ a aplicação definida por $F(x, y, z) = (x + y + z, 2x - 3y + 4z)$. Mostre que F é linear.

Vamos argumentar com matrizes. Escrevendo vetores como colunas, a aplicação F pode ser escrita na forma $F(v) = A v$, onde $v = [x, y, z]^T$ e

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -3 & 4 \end{bmatrix}$$

Então, usando propriedades de matrizes, obtemos

$$F(v + w) = A(v + w) = Av + Aw = F(v) + F(w)$$

e

$$F(kv) = A(kv) = k(Av) = kF(v)$$

Assim, F é linear.

5.11 Mostre que não são lineares as aplicações dadas.

(a) $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (xy, x)$.

(b) $F : \mathbf{R}^2 \to \mathbf{R}^3$ definida por $F(x, y) = (x + 3,\ 2y,\ x + y)$.

(c) $F : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (|x|,\ y + z)$.

(a) Sejam $v = (1, 2)$ e $w = (3, 4)$. Então $v + w = (4, 6)$, bem como

$$F(v) = (1(2), 1) = (2, 1) \qquad \text{e} \qquad F(w) = (3(4), 3) = (12, 3)$$

Logo,

$$F(v + w) = (4(6), 4) = (24, 6) \neq F(v) + F(w)$$

(b) Como $F(0, 0) = (3, 0, 0) \neq (0, 0, 0)$, F não pode ser linear.

(c) Sejam $v = (1, 2, 3)$ e $k = -3$. Então $kv = (-3, -6, -9)$. Temos

$$F(v) = (1, 5) \qquad \text{e} \qquad kF(v) = -3(1, 5) = (-3, -15).$$

Assim,

$$F(kv) = F(-3, -6, -9) = (3, -15) \neq kF(v)$$

Por isso, F não é linear.

5.12 Seja V o espaço vetorial das matrizes reais quadradas de ordem n. Seja M uma matriz arbitrária fixada de V. Seja $F : V \to V$ definida por $F(A) = AM + MA$, para cada matriz A de V. Mostre que F é linear.

Dadas matrizes A e B de V quaisquer e um escalar k qualquer, temos

$$\begin{aligned} F(A + B) &= (A + B)M + M(A + B) = AM + BM + MA + MB \\ &= (AM + MA) = (BM + MB) = F(A) + F(B) \end{aligned}$$

e

$$F(kA) = (kA)M + M(kA) = k(AM) + k(MA) = k(AM + MA) = kF(A)$$

Assim, F é linear.

5.13 Demonstre o Teorema 5.2. Sejam V e U espaços vetoriais sobre um corpo K. Sejam $\{v_1, v_2, ..., v_n\}$ uma base de V e $u_1, u_2, ..., u_n$ vetores quaisquer de U. Então existe uma única transformação linear $F : V \to U$ tal que $F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

A demonstração será dividida em três etapas. (1) Definimos a aplicação $F : V \to U$ de tal modo que $F(v_i) = u_i, i = 1, \ldots, n$. (2) Mostramos que F é linear. (3) Mostramos a unicidade de F.

Passo 1 Seja $v \in V$. Como $\{v_1, v_2, ..., v_n\}$ é uma base de V, existem escalares únicos $a_1, \ldots, a_n \in K$ para os quais $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$. Definimos $F : V \to U$ por

$$F(v) = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

(Pela unicidade dos a_i, a aplicação F está bem definida.) Agora, para $i = 1, ..., n$,

$$v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_n$$

Logo,

$$F(v_i) = 0u_1 + \cdots + 1u_i + \cdots + 0u_n = u_i$$

Assim, completamos o primeiro passo da demonstração.

Passo 2 Sejam dados $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ e $w = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Então

$$v + w = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \cdots + (a_n + b_n)v_n$$

e, dado um escalar $k \in K$ qualquer, $kv = ka_1v_1 + ka_2v_2 + \cdots + ka_nv_n$. Pela definição da aplicação F,

$$F(v) = a_1u_1 + a_2u_2 + \cdots + a_nv_n \quad \text{e} \quad F(w) = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

Logo,

$$\begin{aligned} F(v + w) &= (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_n + b_n)u_n \\ &= (a_1u_1 + a_2u_2 + \cdots + a_nu_n) + (b_1u_1 + b_2u_2 + \cdots + b_nu_n) \\ &= F(v) + F(w) \end{aligned}$$

e

$$F(kv) = k(a_1u_1 + a_2u_2 + \cdots + a_nu_n) = kF(v)$$

Assim, F é linear.

Passo 3 Seja $G: V \to U$ uma transformação linear tal que $G(v_1) = u_i, i = 1, \ldots, n$. Seja

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

Então

$$\begin{aligned} G(v) &= G(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1G(v_1) + a_2G(v_2) + \cdots + a_nG(v_n) \\ &= a_1u_1 + a_2u_2 + \cdots + a_nu_n = F(v) \end{aligned}$$

Como $G(v) = F(v)$, para cada $v \in V$, resulta $G = F$. Assim, F é única e o teorema está demonstrado.

5.14 Seja $F: \mathbf{R}^2 \to \mathbf{R}^2$ a transformação linear tal que $F(1, 2) = (2, 3)$ e $F(0, 1) = (1, 4)$. [Observe que $\{(1, 2), (0, 1)\}$ é uma base de $\mathbf{R}^2$, de modo que uma tal transformação linear F existe e é única, pelo Teorema 5.2.] Encontre uma fórmula para F, ou seja, obtenha $F(a, b)$.

Escrevemos (a, b) como uma combinação linear de $(1, 2)$ e $(0, 1)$ usando incógnitas x e y.

$$(a, b) = x(1, 2) + y(0, 1) = (x, \; 2x + y), \quad \text{portanto,} \quad a = x, b = 2x + y$$

Resolvemos x e y em termos de a e b, para obter $x = a$, $y = -2a + b$. Então

$$F(a, b) = xF(1, 2) + yF(0, 1) = a(2, 3) + (-2a + b)(1, 4) = (b, \; -5a + 4b)$$

5.15 Seja $F: V \to U$ uma transformação linear que é injetora e sobre. Mostre que a aplicação inversa $F^{-1}: U \to V$ também é linear.

Sejam $u, u' \in U$. Como F é injetora e sobre, existem vetores únicos $v, v' \in V$ para os quais $F(v) = u$ e $F(v') = u'$. Como F é linear, também temos

$$F(v + v') = F(v) + F(v') = u + u' \quad \text{e} \quad F(rv) = rF(v) = ru$$

Pela definição de aplicação inversa,

$$F^{-1}(u) = v, \; F^{-1}(u') = v', \; F^{-1}(u + u') = v + v', F^{-1}(ru) = rv$$

Então

$$F^{-1}(u + u') = v + v' = F^{-1}(u) + F^{-1}(u') \quad \text{e} \quad F^{-1}(ru) = rv = rF^{-1}(u)$$

Assim, F^{-1} é linear.

Núcleo e imagem de transformações lineares

5.16 Seja $F : \mathbf{R}^4 \to \mathbf{R}^3$ a transformação linear definida por

$$F(x, y, z, t) = (x - y + z + t, \quad x + 2z - t, \quad x + y + 3z - 3t)$$

Encontre uma base e a dimensão (a) da imagem de F, (b) do núcleo de F.

(a) Calculamos as imagens dos vetores da base canônica de $\mathbf{R}^4$.

$$F(1,0,0,0) = (1,1,1), \qquad F(0,0,1,0) = (1,2,3)$$
$$F(0,1,0,0) = (-1,0,1), \qquad F(0,0,0,1) = (1,-1,-3)$$

Pela Proposição 5.4, os vetores imagem geram Im F. Logo, formamos a matriz cujas linhas são esses vetores imagem e a reduzimos à forma escalonada.

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, (1, 1, 1) e (0, 1, 2) formam uma base de Im F e, portanto, dim(Im F) = 2.

(b) Tomamos $F(v) = 0$, com $v = (x, y, z, t)$, ou seja, tomamos

$$F(x, y, z, t) = (x - y + z + t, \ x + 2z - t, \ x + y + 3z - 3t) = (0, 0, 0)$$

Igualando as entradas correspondentes, obtemos o sistema homogêneo cujo espaço solução é Nuc F.

$$\begin{aligned} x - y + z + t &= 0 \\ x \quad + 2z - t &= 0 \\ x + y + 3z - 3t &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + z + t &= 0 \\ y + z - 2t &= 0 \\ 2y + 2z - 4t &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + z + t &= 0 \\ y + z - 2t &= 0 \end{aligned}$$

As variáveis livres são z e t. Logo, dim(Nuc F) = 2.

(i) Tomando $z = -1$, $t = 0$ obtemos a solução (2, 1, –1, 0).
(ii) Tomando $z = 0$, $t = 1$ obtemos a solução (1, 2, 0, 1).

Assim, (2, 1, –1, 0) e (1, 2, 0, 1) formam uma base de Nuc F.
[Como era de se esperar, dim(Im F) + dim(Nuc F) = 2 + 2 = 4 = dim $\mathbf{R}^4$, o domínio de F.]

5.17 Seja $G : \mathbf{R}^3 \to \mathbf{R}^3$ a transformação linear definida por

$$G(x, y, z) = (x + 2y - z, \quad y + z, \quad x + y - 2z)$$

Encontre uma base e a dimensão (a) da imagem de G, (b) do núcleo de G.

(a) Calculamos as imagens dos vetores da base canônica de $\mathbf{R}^3$.

$$G(1,0,0) = (1,0,1), \qquad G(0,1,0) = (2,1,1), \qquad G(0,0,1) = (-1,1,-2)$$

Pela Proposição 5.4, os vetores imagem geram Im G. Logo, formamos a matriz M cujas linhas são esses vetores imagem e a reduzimos à forma escalonada.

$$M = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, (1, 0, 1) e (0, 1, –1) formam uma base de Im G e, portanto, dim(Im G) = 2.

(b) Tomamos $G(v) = 0$, com $v = (x, y, z)$, ou seja, tomamos

$$G(x, y, z) = (x + 2y - z, \quad y + z, \quad x + y - 2z) = (0, 0, 0)$$

Igualando as entradas correspondentes, obtemos o sistema homogêneo cujo espaço solução é Nuc G.

$$\begin{aligned} x+2y-\ \ z&=0\\ y+\ \ z&=0\\ x+\ \ y-2z&=0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x+2y-z&=0\\ y+z&=0\\ -y-z&=0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x+2y-z&=0\\ y+z&=0 \end{aligned}$$

A única variável livre é z. Logo, dim(Nuc G) = 1. Tomando $z = 1$, obtemos $y = -1$ e $x = 3$. Assim, (3, –1, 1) forma uma base de Nuc G. [Como era de se esperar, dim(Im G) + dim(Nuc G) = 2 + 1 = 3 = dim $\mathbf{R}^3$, o domínio de G.]

5.18 Considere a transformação matricial $A:\mathbf{R}^4 \to \mathbf{R}^3$, com $A=\begin{bmatrix}1&2&3&1\\1&3&5&-2\\3&8&13&-3\end{bmatrix}$. Encontre uma base e a dimensão (a) da imagem de A, (b) do núcleo de A.

(a) O espaço coluna de A é igual a Im A. Logo, reduzimos A^T à forma escalonada.

$$A^T=\begin{bmatrix}1&1&3\\2&3&8\\3&5&13\\1&-2&-3\end{bmatrix}\sim\begin{bmatrix}1&1&3\\0&1&2\\0&2&4\\0&-3&-6\end{bmatrix}\sim\begin{bmatrix}1&1&3\\0&1&2\\0&0&0\\0&0&0\end{bmatrix}$$

Assim, {(1, 1, 3), (0, 1, 2)} é uma base de Im A e, portanto, dim(Im A) = 2.

(b) Sabemos que Nuc A é o espaço solução do sistema homogêneo $AX = 0$, com $X = (x, y, z, t)^T$. Logo, reduzimos a matriz A de coeficientes à forma escalonada.

$$\begin{bmatrix}1&2&3&1\\0&1&2&-3\\0&2&4&-6\end{bmatrix}\sim\begin{bmatrix}1&2&3&1\\0&1&2&-3\\0&0&0&0\end{bmatrix} \quad \text{ou} \quad \begin{aligned} x+2y+3z+\ \ t&=0\\ y+2z-3t&=0 \end{aligned}$$

As variáveis livres são z e t. Logo, dim(Nuc A) = 2.

(i) Tomando $z = 1$, $t = 0$ obtemos a solução (1, –2, 1, 0).
(ii) Tomando $z = 0$, $t = 1$ obtemos a solução (–7, 3, 0, 1).

Assim, (1, –2, 1, 0) e (–7, 3, 0, 1) formam uma base de Nuc A.

5.19 Encontre uma transformação linear $F:\mathbf{R}^3 \to \mathbf{R}^4$ cuja imagem seja gerada por (1, 2, 0, –4) e (2, 0, –1, –3).

Formamos a matriz 4×3 cujas colunas consistem somente nos vetores dados, digamos,

$$A=\begin{bmatrix}1&2&2\\2&0&0\\0&-1&-1\\-4&-3&-3\end{bmatrix}$$

Como A determina uma transformação linear $A:\mathbf{R}^3 \to \mathbf{R}^4$ cuja imagem é gerada pelas colunas de A, vemos que a transformação matricial A satisfaz o exigido.

5.20 Suponha que $f: V \to U$ seja linear com núcleo W e que $f(v) = u$. Mostre que a "classe lateral" $v + W = \{v + w : w \in W\}$ é a pré-imagem de u, ou seja, $f^{-1}(u) = v + W$.

Precisamos provar que (i) $f^{-1}(u) \subseteq v + W$ e (ii) $v + W \subseteq f^{-1}(u)$.
Começamos com (i). Seja $v' \in f^{-1}(u)$. Então $f(v') = u$, de modo que

$$f(v'-v)=f(v')-f(v)=u-u=0$$

ou seja, $v' - v \in W$. Assim, $v' = v + (v' - v) \in v + W$ e, portanto, e $f^{-1}(u) \subseteq v + W$.

Agora provamos (ii). Seja $v' \in v + W$. Então $v' = v + w$, com $w \in W$. Como W é o núcleo de f, temos $f(w) = 0$. Dessa forma,

$$f(v') = f(v + w) + f(v) + f(w) = f(v) + 0 = f(v) = u$$

Assim, $v' \in f^{-1}(u)$ e, portanto, $v + W \subseteq f^{-1}(u)$.

Essas duas inclusões implicam $f^{-1}(u) = v + W$.

5.21 Sejam $F : V \to U$ e $G : U \to W$ lineares. Prove que

(a) $\text{pos}(G \circ F) \leq \text{pos}(G)$, (b) $\text{pos}(G \circ F) \leq \text{pos}(F)$

(a) Como $F(V) \subseteq U$, também temos $G(F(V)) \subseteq G(U)$ e, portanto, $\dim[G(F(V))] \leq \dim[G(U)]$. Então $\text{pos}(G \circ F) = \dim[(G \circ F)(V)] = \dim[G(F(V))] \leq \dim[G(U)] = \text{pos}(G)$.

(b) Temos $\dim[G(F(V))] \leq \dim[F(V)]$. Logo,

$$\text{pos}(G \circ F) = \dim[(G \circ F)(V)] = \dim[G(F(V))] \leq \dim[F(V)] = \text{pos}(F)$$

5.22 Demonstre o Teorema 5.3. Seja $F : V \to U$ uma transformação linear. Então

(a) Im F é um subespaço de U, (b) Nuc F é um subespaço de V.

(a) Como $F(0) = 0$, temos $0 \in \text{Im } F$. Sejam, agora, $u, u' \in \text{Im } F$ e $a, b \in K$. Como u e u' pertencem à imagem de F, existem vetores $v, v' \in V$ tais que $F(v) = u$ e $F(v') = u'$. Então

$$F(av + bv') = aF(v) + bF(v') = au + bu' \in \text{Im } F$$

Assim, a imagem de F é um subespaço de U.

(b) Como $F(0) = 0$, temos $0 \in \text{Nuc } F$. Sejam, agora, $v, w \in \text{Ker } F$ e $a, b \in K$. Como v e w pertencem ao núcleo de F, $F(v) = 0$ e $F(w) = 0$. Assim,

$$F(av + bw) = aF(v) + bF(w) = a0 + b0 = 0 + 0 = 0, \qquad \text{e, portanto,} \qquad av + bw \in \text{Nuc } F$$

Assim, o núcleo de F é um subespaço de V.

5.23 Demonstre o Teorema 5.6. Sejam V um espaço de dimensão finita e $F : V \to U$ uma transformação linear. Então

$$\dim V = \dim(\text{Nuc } F) + \dim(\text{Im } F) = \text{nul}(F) + \text{pos}(F)$$

Suponha que $\dim(\text{Nuc } F) = r$, que $\{w_1, ..., w_r\}$ seja uma base de Nuc F, que $\dim(\text{Im } F) = s$ e que $\{u_1, ..., u_s\}$ seja uma base de Im F. (Pela Proposição 5.4, Im F tem dimensão finita.) Como cada $u_j \in \text{Im } F$, existem vetores $v_1, ..., v_s$ de V tais que $F(v_1) = u_1, \dots, F(v_s) = u_s$. Afirmamos que o conjunto

$$B = \{w_1, \dots, w_r, v_1, \dots, v_s\}$$

é uma base de V, ou seja, que (i) B gera V e (ii) B é linearmente independente. Uma vez demonstrados (i) e (ii), resulta que $\dim V = r + s = \dim(\text{Nuc } F) + \dim(\text{Im } F)$.

(i) *B gera V*. Seja $v \in V$. Então $F(v) \in \text{Im } F$. Como os u_j geram Im F, existem escalares $a_1, ..., a_s$ tais que $F(v) = a_1u_1 + \cdots + a_su_s$. Denotemos $\hat{v} = a_1v_1 + \cdots + a_sv_s - v$. Então

$$\begin{aligned} F(\hat{v}) &= F(a_1v_1 + \cdots + a_sv_s - v) = a_1F(v_1) + \cdots + a_sF(v_s) - F(v) \\ &= a_1u_1 + \cdots + a_su_s - F(v) = 0 \end{aligned}$$

Assim, $\hat{v} \in \text{Nuc } F$. Como os W_i; geram Nuc F, existem escalares $b_1, ..., b_r$ tais que

$$\hat{v} = b_1w_1 + \cdots + b_rw_r = a_1v_1 + \cdots + a_sv_s - v$$

Em vista disso,

$$v = a_1v_1 + \cdots + a_sv_s - b_1w_1 - \cdots - b_rw_r$$

Assim, B gera V.

(ii) *B é linearmente independente*. Suponha que

$$x_1w_1 + \cdots + x_rw_r + y_1v_1 + \cdots + y_sv_s = 0 \tag{1}$$

com $x_i, y_j \in K$. Então

$$\begin{aligned} 0 = F(0) &= F(x_1w_1 + \cdots + x_rw_r + y_1v_1 + \cdots + y_sv_s) \\ &= x_1F(w_1) + \cdots + x_rF(w_r) + y_1F(v_1) + \cdots + y_sF(v_s) \end{aligned} \tag{2}$$

Mas $F(w_i) = 0$, pois $w_i \in \text{Nuc } F$, e $F(v_j) = u_j$. Substituindo em (2), obtemos $y_1u_1 + \cdots + y_su_s = 0$. Como os u_j são linearmente independentes, cada $y_j = 0$. Substituindo em (1), obtemos $x_1w_1 + \cdots + x_rw_r = 0$. Como os w_i são linearmente independentes, cada $x_i = 0$. Assim, B é linearmente independente.

Transformações lineares singulares e não singulares, isomorfismos

5.24 Decida se a transformação linear dada é não singular. Se não for, encontre um vetor v não nulo cuja imagem seja 0.

(a) $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (x - y,\ x - 2y)$.

(b) $G : \mathbf{R}^2 \to \mathbf{R}^2$, definida por $G(x, y) = (2x - 4y,\ 3x - 6y)$.

(a) Encontramos Nuc F tomando $F(v) = 0$, com $v = (x, y)$, obtendo

$$(x - y,\ x - 2y) = (0, 0) \quad \text{ou} \quad \begin{aligned} x - \ \ y &= 0 \\ x - 2y &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y &= 0 \\ -y &= 0 \end{aligned}$$

A única solução é $x = 0, y = 0$. Assim, F é não singular.

(b) Tomamos $G(x, y) = (0, 0)$ para encontrar Nuc G.

$$(2x - 4y,\ 3x - 6y) = (0, 0) \quad \text{ou} \quad \begin{aligned} 2x - 4y &= 0 \\ 3x - 6y &= 0 \end{aligned} \quad \text{ou} \quad x - 2y = 0$$

O sistema tem soluções não nulas, porque y é uma variável livre. Logo, G é singular. Tomando $y = 1$, obtemos a solução $v = (2, 1)$, que é um vetor não nulo tal que $G(v) = 0$.

5.25 A transformação linear $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (x - y,\ x - 2y)$ é não singular, segundo o Problema 5.24 precedente. Encontre uma fórmula para F^{-1}.

Tomando $F(x, y) = (a, b)$, temos $F^{-1}(a, b) = (x, y)$. Calculando,

$$(x - y,\ x - 2y) = (a, b) \quad \text{ou} \quad \begin{aligned} x - \ \ y &= a \\ x - 2y &= b \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y &= a \\ y &= a - b \end{aligned}$$

Resolvendo x e y em termos de a e b, obtemos $x = 2a - b$, $y = a - b$. Assim,

$$F^{-1}(a, b) = (2a - b,\ a - b) \quad \text{ou} \quad F^{-1}(x, y) = (2x - y,\ x - y)$$

(A segunda equação foi obtida substituindo a e b por x e y, respectivamente.)

5.26 Seja $G : \mathbf{R}^2 \to \mathbf{R}^3$ definida por $G(x, y) = (x + y,\ x - 2y,\ 3x + y)$.

(a) Mostre que G é não singular. (b) Encontre uma fórmula para G^{-1}.

(a) Tomamos $G(x, y) = (0, 0, 0)$ para encontrar Nuc G. Temos

$$(x + y,\ x - 2y,\ 3x + y) = (0, 0, 0) \quad \text{ou} \quad x + y = 0,\ x - 2y = 0,\ 3x + y = 0$$

A única solução é $x = 0, y = 0$, portanto, G é não singular.

(b) Embora G seja não singular, não é invertível, porque $\mathbf{R}^2$ e $\mathbf{R}^3$ têm dimensões diferentes. (Assim, não podemos aplicar o Teorema 5.9.) Em vista disso, não existe G^{-1}.

5.27 Suponha que $F : V \to U$ seja linear com V seja de dimensão finita. Mostre que V e a imagem de F têm a mesma dimensão se, e só se, F é não singular. Encontre todas as transformações lineares não singulares $T : \mathbf{R}^4 \to \mathbf{R}^3$.

Pelo Teorema 5.6, dim V = dim(Im F) + dim(Nuc F). Logo, V e Im F têm a mesma dimensão se, e somente se, dim(Nuc F) = 0, ou Nuc $F = \{0\}$ (ou seja, se, e só se, F é não singular).
Como dim $\mathbf{R}^3$ é menor do que dim $\mathbf{R}^4$, temos que dim(Im T) é menor do que a dimensão do domínio $\mathbf{R}^4$ de T. Em vista disso, nenhuma transformação linear $T : \mathbf{R}^4 \to \mathbf{R}^3$ pode ser não singular.

5.28 Demonstre o Teorema 5.7. Seja $F : V \to U$ uma transformação linear não singular. Então a imagem de qualquer conjunto linearmente independente é linearmente independente.

Suponha que $v_1, v_2, ..., v_n$ sejam vetores linearmente independentes de V. Afirmamos que $F(v_1), F(v_2), ..., F(v_n)$ também são linearmente independentes. Suponha que $a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n) = 0$, com $a_i \in K$. Como F é linear, $F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = 0$. Logo,

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n \in \text{Nuc } F$$

Mas F é não singular, ou seja, Nuc $F = \{0\}$. Logo, $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Como os v_i são linearmente independentes, todos os a_i são nulos. Por isso, os $F(v_i)$ são linearmente independentes. Assim, demonstramos o teorema.

5.29 Demonstre o Teorema 5.9. Seja V um espaço vetorial de dimensão finita e suponha que dim V = dim U. Seja $F : V \to U$ linear. Então F é um isomorfismo se, e só se, F é não singular.

Se F é um isomorfismo, então somente 0 é levado em 0, de modo que F é não singular. Reciprocamente, suponha que F seja não singular. Então dim(Nuc F) = 0. Pelo Teorema 5.6, dim V = dim(Nuc F) + dim(Im F). Assim,

$$\dim U = \dim V = \dim(\text{Im } F)$$

Como U tem dimensão finita, Im $F = U$. Isso significa que F leva V sobre U. Assim, F é injetora e sobre, ou seja, um isomorfismo.

Operações com transformações lineares

5.30 Defina $F : \mathbf{R}^3 \to \mathbf{R}^2$ e $G : \mathbf{R}^3 \to \mathbf{R}^2$ por $F(x, y, z) = (2x,\ y + z)$ e $G(x, y, z) = (x - z,\ y)$. Encontre fórmulas que definam as transformações (a) $F + G$, (b) $3F$, (c) $2F - 5G$.

(a) $(F + G)(x, y, z) = F(x, y, z) + G(x, y, z) = (2x,\ y + z) + (x - z,\ y) = (3x - z,\ 2y + z)$

(b) $(3F)(x, y, z) = 3F(x, y, z) = 3(2x,\ y + z) = (6x,\ 3y + 3z)$

(c) $(2F - 5G)(x, y, z) = 2F(x, y, z) - 5G(x, y, z) = 2(2x,\ y + z) - 5(x - z,\ y)$
$= (4x,\ 2y + 2z) + (-5x + 5z,\ -5y) = (-x + 5z,\ -3y + 2z)$

5.31 Sejam $F : \mathbf{R}^3 \to \mathbf{R}^2$ e $G : \mathbf{R}^2 \to \mathbf{R}^2$ definidas por $F(x, y, z) = (2x,\ y + z)$ e $G(x, y) = (y, x)$. Deduza fórmulas que definam as transformações (a) $G \circ F$, (b) $F \circ G$.

(a) $(G \circ F)(x, y, z) = G(F(x, y, z)) = G(2x,\ y + z) = (y + z,\ 2x)$

(b) A aplicação $F \circ G$ não está definida, porque a imagem de G não está contida no domínio de F.

5.32 Demonstre. (a) A transformação nula 0, definida por $0(v) = 0 \in U$, para cada $v \in V$, é o elemento nulo de Hom(V, U). (b) O negativo de $F \in \text{Hom}(V, U)$ é a transformação $(-1)F$, ou seja, $-F = (-1)F$.

Seja $F \in \text{Hom}(V, U)$. Então, para cada $v \in V$,

(a) $$(F + 0)(v) = F(v) + 0(v) = F(v) + 0 = F(v)$$

Como $(F + 0)(v) = F(v)$, para cada $v \in V$, temos $F + 0 = F$. Analogamente, $0 + F = F$.

(b) $$(F + (-1)F)(v) = F(v) + (-1)F(v) = F(v) - F(v) = 0 = 0(v)$$

Logo, $F + (-1)F = 0$. Analogamente, $(-1)F + F = 0$. Assim, $-F = (-1)F$.

5.33 Suponha que $F_1, F_2, ..., F_n$ sejam transformações lineares de V em U. Mostre que, dados escalares $a_1, a_2, ..., a_n$ quaisquer e um vetor $v \in V$ qualquer,

$$(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) = a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)$$

A transformação a_1F_1 é definida por $(a_1F_1)(v) = a_1F_1(v)$. Logo, a afirmação é válida para $n = 1$. Dessa forma, por indução, obtemos

$$\begin{aligned}(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) &= (a_1F_1)(v) + (a_2F_2 + \cdots + a_nF_n)(v) \\ &= a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)\end{aligned}$$

5.34 Considere as transformações lineares $F : \mathbf{R}^3 \to \mathbf{R}^2$, $G : \mathbf{R}^3 \to \mathbf{R}^2$, $H : \mathbf{R}^3 \to \mathbf{R}^2$ definidas por

$F(x, y, z) = (x + y + z,\ x + y)$, $G(x, y, z) = (2x + z,\ x + y)$, $H(x, y, z) = (2y,\ x)$

Mostre que F, G, H são linearmente independentes [como elementos de $\text{Hom}(\mathbf{R}^3, \mathbf{R}^2)$].

Dados escalares $a, b, c \in K$, suponha que

$$aF + bG + cH = 0 \tag{1}$$

(Aqui, 0 é a transformação nula.) Para $e_1 = (1, 0, 0) \in \mathbf{R}^3$, temos $0(e_1) = (0, 0)$ e

$$\begin{aligned}(aF + bG + cH)(e_1) &= aF(1, 0, 0) + bG(1, 0, 0) + cH(1, 0, 0) \\ &= a(1, 1) + b(2, 1) + c(0, 1) = (a + 2b,\quad a + b + c)\end{aligned}$$

Assim, por (1), $(a + 2b, a + b + c) = (0, 0)$ e, portanto,

$$a + 2b = 0 \qquad \text{e} \qquad a + b + c = 0 \tag{2}$$

Analogamente, para $e_2 = (0, 1, 0) \in \mathbf{R}^3$, temos $0(e_2) = (0, 0)$ e

$$\begin{aligned}(aF + bG + cH)(e_2) &= aF(0, 1, 0) + bG(0, 1, 0) + cH(0, 1, 0) \\ &= a(1, 1) + b(0, 1) + c(2, 0) = (a + 2c,\quad a + b)\end{aligned}$$

Assim,

$$a + 2c = 0 \qquad \text{e} \qquad a + b = 0 \tag{3}$$

Usando (2) e (3), obtemos

$$a = 0, b = 0, c = 0 \tag{4}$$

Como (1) implica (4), as transformações F, G, H são linearmente independentes.

5.35 Seja k um escalar não nulo. Mostre que uma transformação linear T é singular se, e só se, kT é singular. Em particular, T é singular se, e só se, $-T$ é singular.

Suponha que T seja singular. Então $T(v) = 0$, para algum vetor $v \neq 0$. Logo,

$$(kT)(v) = kT(v) = k0 = 0$$

e, portanto, kT é singular.

Agora, suponha que kT seja singular. Então $(kT)(w) = 0$, para algum vetor $w \neq 0$. Logo,

$$T(kw) = kT(w) = (kT)(w) = 0$$

Mas, $k \neq 0$ e $w \neq 0$ implicam $kw \neq 0$. Assim, T é singular.

5.36 Encontre a dimensão d de

(a) $\text{Hom}(\mathbf{R}^3, \mathbf{R}^4)$, (b) $\text{Hom}(\mathbf{R}^5, \mathbf{R}^3)$, (c) $\text{Hom}(\mathbf{P}_3(t), \mathbf{R}^2)$, (d) $\text{Hom}(\mathbf{M}_{2,3}, \mathbf{R}^4)$.

Usamos que $\dim[\text{Hom}(V, U)] = mn$ se $\dim V = m$ e $\dim U = n$.

(a) $d = 3(4) = 12$.

(b) $d = 5(3) = 15$.

(c) Como $\dim \mathbf{P}_3(t) = 4$, $d = 4(2) = 8$.

(d) Como $\dim \mathbf{M}_{2,3} = 6$, $d = 6(4) = 24$.

5.37 Demonstre o Teorema 5.11. Se dim $V = m$ e dim $U = n$, então dim[Hom(V, U)] = mn.

Sejam $\{v_1, ..., v_m\}$ uma base de V e $\{u_1, ..., u_n\}$ uma base de U. Pelo Teorema 5.2, cada transformação linear de Hom(V, U) fica univocamente determinada associando valores arbitrários de U aos elementos v_i da base de V. Definimos

$$F_{ij} \in \text{Hom}(V, U), \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n$$

como sendo a transformação linear dada por $F_{ij}(v_i) = u_j$ e $F_{ij}(v_k) = 0$, com $k \neq i$. Ou seja, F_{ij} leva v_i em u_j e os demais v em 0. Observe que $\{F_{ij}\}$ contém exatamente mn elementos. Dessa forma, o teorema estará demonstrado se mostrarmos que esse conjunto é uma base de Hom(V, U).

Demonstremos que $\{F_{ij}\}$ *gera* Hom(V, U). Considere uma transformação $F \in \text{Hom}(V, U)$ arbitrária. Sejam $F(v_1) = w_1$, $F(v_2) = w_2$, ..., $F(v_m) = w_m$. Como $w_k \in U$, cada w_k é uma combinação linear dos u_i, digamos,

$$w_k = a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n, \qquad k = 1, \ldots, m, \quad a_{ij} \in K \tag{1}$$

Considere a transformação linear $G = \sum_{i=1}^m \sum_{j=1}^n a_{ij}F_{ij}$. Como G é uma combinação linear dos F_{ij}, a demonstração de que $\{F_{ij}\}$ gera Hom(V, U) termina se mostrarmos que $F = G$.
Para isso, calculamos $G(v_k)$, para $k = 1, ..., m$. Como $F_{ij}(v_k) = 0$, para $k \neq i$, e $F_{ki}(v_k) = u_i$, temos

$$\begin{aligned} G(v_k) &= \sum_{i=1}^m \sum_{j=1}^n a_{ij}F_{ij}(v_k) = \sum_{j=1}^n a_{kj}F_{kj}(v_k) = \sum_{j=1}^n a_{kj}u_j \\ &= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n \end{aligned}$$

Assim, por (1), $G(v_k) = w_k$, para cada k. Mas $F(v_k) = w_k$, para cada k. Em vista disso, pelo Teorema 5.2, $F = G$. Assim, $\{F_{ij}\}$ gera Hom(V, U).
Demonstremos que $\{F_{ij}\}$ *é linearmente independente*. Considere escalares $c_{ij} \in K$ tais que

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij}F_{ij} = 0$$

Para cada v_k, com $k = 1, ..., m$,

$$\begin{aligned} 0 = 0(v_k) &= \sum_{i=1}^m \sum_{j=1}^n c_{ij}F_{ij}(v_k) = \sum_{j=1}^n c_{kj}F_{kj}(v_k) = \sum_{j=1}^n c_{kj}u_j \\ &= c_{k1}u_1 + c_{k2}u_2 + \cdots + c_{kn}u_n \end{aligned}$$

Mas os u_i são linearmente independentes, portanto, para $k = 1, ..., m$, temos $c_{k1} = 0, c_{k2} = 0, \ldots, c_{kn} = 0$. Em outras palavras, todos os $c_{ij} = 0$ e, portanto, $\{F_{ij}\}$ é linearmente independente.

5.38 Demonstre o Teorema 5.12. (i) $G \circ (F + F') = G \circ F + G \circ F'$. (ii) $(G + G') \circ F = G \circ F + G' \circ F$. (iii) $k(G \circ F) = (kG) \circ F = G \circ (kF)$.

(i) Dado $v \in V$ qualquer,

$$\begin{aligned} (G \circ (F + F'))(v) &= G((F + F')(v)) = G(F(v) + F'(v)) \\ &= G(F(v)) + G(F'(v)) = (G \circ F)(v) + (G \circ F')(v) = (G \circ F + G \circ F')(v) \end{aligned}$$

Assim, $G \circ (F + F') = G \circ F + G \circ F'$.

(ii) Dado $v \in V$ qualquer,

$$\begin{aligned} ((G + G') \circ F)(v) &= (G + G')(F(v)) = G(F(v)) + G'(F(v)) \\ &= (G \circ F)(v) + (G' \circ F)(v) = (G \circ F + G' \circ F)(v) \end{aligned}$$

Assim, $(G + G') \circ F = G \circ F + G' \circ F$.

(iii) Dado $v \in V$ qualquer,

$$(k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = (kG)(F(v)) = (kG \circ F)(v)$$

e

$$(k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = G(kF(v)) = G((kF)(v)) = (G \circ kF)(v)$$

Por isso, $k(G \circ F) = (kG) \circ F = G \circ (kF)$. (Enfatizamos que a maneira de demonstrar que duas aplicações são iguais consiste em mostrar que cada uma delas associa a mesma imagem a cada ponto do domínio.)

Álgebra das transformações lineares

5.39 Sejam F e G os operadores lineares de $\mathbf{R}^2$ definidos por $F(x, y) = (y, x)$ e $G(x, y) = (0, x)$. Encontre fórmulas que definam os operadores dados.

(a) $F + G$, (b) $2F - 3G$, (c) FG, (d) GF, (e) F^2, (f) G^2.

(a) $(F + G)(x, y) = F(x, y) + G(x, y) = (y, x) + (0, x) = (y, 2x)$.

(b) $(2F - 3G)(x, y) = 2F(x, y) - 3G(x, y) = 2(y, x) - 3(0, x) = (2y, -x)$.

(c) $(FG)(x, y) = F(G(x, y)) = F(0, x) = (x, 0)$.

(d) $(GF)(x, y) = G(F(x, y)) = G(y, x) = (0, y)$.

(e) $F^2(x, y) = F(F(x, y)) = F(y, x) = (x, y)$. (Observe que $F^2 = I$, o operador identidade.)

(f) $G^2(x, y) = G(G(x, y)) = G(0, x) = (0, 0)$. (Observe que $G^2 = 0$, o operador nulo.)

5.40 Considere o operador T de $\mathbf{R}^3$ definido por $T(x, y, z) = (2x, 4x - y, 2x + 3y - z)$. (a) Mostre que T é invertível. Encontre fórmula para (b) T^{-1}, (c) T^2, (d) T^{-2}.

(a) Seja $W = \text{Nuc}\, T$. Basta mostrar que T é não singular (ou seja, que $W = \{0\}$.) Tomando $T(x, y, z) = (0, 0, 0)$, obtemos

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (0, 0, 0)$$

Assim, W é o espaço solução do sistema homogêneo

$$2x = 0, \qquad 4x - y = 0, \qquad 2x + 3y - z = 0$$

cuja única solução é a trivial (0, 0, 0). Logo, $W = \{0\}$. Assim, T é não singular e, portanto, invertível.

(b) Tomemos $T(x, y, z) = (r, s, t)$ [e, portanto, $T^{-1}(r, s, t) = (x, y, z)$.] Temos

$$(2x,\ 4x - y,\ 2x + 3y - z) = (r, s, t) \qquad \text{ou} \qquad 2x = r, \quad 4x - y = s, \quad 2x + 3y - z = t$$

Resolvendo x, y, z em termos de r, s, t, obtemos $x = \frac{1}{2}r$, $y = 2r - s$, $z = 7r - 3s - t$. Assim,

$$T^{-1}(r, s, t) = (\tfrac{1}{2}r,\ 2r - s,\ 7r - 3s - t) \qquad \text{ou} \qquad T^{-1}(x, y, z) = (\tfrac{1}{2}x,\ 2x - y,\ 7x - 3y - z)$$

(c) Aplicamos T duas vezes para obter

$$\begin{aligned} T^2(x, y, z) &= T(2x,\ 4x - y,\ 2x + 3y - z) \\ &= [4x, \quad 4(2x) - (4x - y), \quad 2(2x) + 3(4x - y) - (2x + 3y - z)] \\ &= (4x,\ 4x + y,\ 14x - 6y + z) \end{aligned}$$

(d) Aplicamos T^{-1} duas vezes para obter

$$\begin{aligned} T^{-2}(x, y, z) &= T^{-2}(\tfrac{1}{2}x,\ 2x - y,\ 7x - 3y - z) \\ &= [\tfrac{1}{4}x, \quad 2(\tfrac{1}{2}x) - (2x - y), \quad 7(\tfrac{1}{2}x) - 3(2x - y) - (7x - 3y - z)] \\ &= (\tfrac{1}{4}x, \quad -x + y, \quad -\tfrac{19}{2}x + 6y + z) \end{aligned}$$

5.41 Sejam V um espaço de dimensão finita e T um operador linear de V para o qual $TR = I$, para algum operador linear R de V. (Dizemos que R é uma *inversa à direita* de T.)

(a) Mostre que T é invertível. (b) Mostre que $R = T^{-1}$.

(c) Dê um exemplo mostrando que as afirmações precedentes não valem se V for de dimensão infinita.

(a) Seja dim $V = n$. Pelo Teorema 5.14, T é invertível se, e só se, T é sobre, de modo que é invertível se, e só se, pos(T) = n. Temos $n = \text{pos}(I) = \text{pos}(TR) \leq \text{pos}(T) \leq n$. Logo, pos($T$) = n e T é invertível.

(b) $TT^{-1} = T^{-1}T = I$. Então $R = IR = (T^{-1}T)R = T^{-1}(TR) = T^{-1}I = T^{-1}$.

(c) Seja V o espaço dos polinômios em t sobre K, digamos, do tipo $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s$. Sejam T e R os operadores de V definidos por

$$T(p(t)) = 0 + a_1 + a_2 t + \cdots + a_s t^{s-1} \quad \text{e} \quad R(p(t)) = a_0 t + a_1 t^2 + \cdots + a_s t^{s+1}$$

Temos

$$(TR)(p(t)) = T(R(p(t))) = T(a_0 t + a_1 t^2 + \cdots + a_s t^{s+1}) = a_0 + a_1 t + \cdots + a_s t^s = p(t)$$

e, portanto, $TR = I$, o operador identidade. Por outro lado, se $k \in K$ e $k \neq 0$, então

$$(RT)(k) = R(T(k)) = R(0) = 0 \neq k$$

Por isso, $RT \neq I$.

5.42 Sejam F e G operadores lineares de $\mathbf{R}^2$ definidos por $F(x, y) = (0, x)$ e $G(x, y) = (x, 0)$. Mostre que (a) $GF = 0$, o operador nulo, mas $FG \neq 0$. (b) $G^2 = G$.

(a) $(GF)(x, y) = G(F(x, y)) = G(0, x) = (0, 0)$. Como GF associa $0 = (0, 0)$ a cada vetor (x,y) $\mathbf{R}^2$, é o operador nulo, ou seja, $GF = 0$.

Por outro lado, $(FG)(x, y) = F(G(x, y)) = F(x, 0) = (0, x)$. Por exemplo, $(FG)(2, 3) = (0, 2)$. Assim, $FG \neq 0$, pois não associa $0 = (0, 0)$ a cada vetor de $\mathbf{R}^2$.

(b) Dado qualquer vetor (x, y) em $\mathbf{R}^2$, temos $G^2(x, y) = G(G(x, y)) = G(x, 0) = (x, 0) = G(x, y)$. Logo, $G^2 = G$.

5.43 Encontre as dimensões de (a) $A(\mathbf{R}^4)$, (b) $A(\mathbf{P}_2(t))$, (c) $A(\mathbf{M}_{2,3})$.

Usamos que $\dim[A(V)] = n^2$ se $\dim V = n$. Logo, (a) $\dim[A(\mathbf{R}^4)] = 4^2 = 16$, (b) $\dim[A(\mathbf{P}_2(t))] = 3^2 = 9$, (c) $\dim[A(\mathbf{M}_{2,3})] = 6^2 = 36$.

5.44 Seja E um operador linear de V para o qual $E^2 = E$. (Tais operadores são denominados *idempotentes*.) Sejam U a imagem e W o núcleo de E. Prove as afirmações seguintes.

(a) Se $u \in U$, então $E(u) = u$ (ou seja, E é a transformação identidade em U).

(b) Se $E \neq I$, então E é singular (ou seja, $E(v) = 0$, para algum $v \neq 0$).

(c) $V = U \oplus W$.

(a) Se $u \in U$, então $E(v) = u$ para algum $v \in V$. Logo, usando $E^2 = E$, obtemos

$$u = E(v) = E^2(v) = E(E(v)) = E(u)$$

(b) Se $E \neq I$, então $E(v) = u$, para algum $v \in V$, $E(v) = u$, com $v \neq u$. Por (a), $E(u) = u$. Assim,

$$E(v - u) = E(v) - E(u) = u - u = 0, \quad \text{com} \quad v - u \neq 0$$

(c) Mostremos que $V = U + W$. Seja $v \in V$. Tomemos $u = E(v)$ e $w = v - E(v)$. Então

$$v = E(v) + v - E(v) = u + w$$

Por definição, $u = E(v) \in U$ a imagem de E. Agora mostramos que $w \in W$, o núcleo de E, como segue.

$$E(w) = E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

Logo, $w \in W$. Assim, $V = U + W$.

Agora mostremos que $U \cap W = \{0\}$. Dado $v \in U \cap W$, temos $E(v) = v$ pela parte (a), pois $v \in U$. Como $v \in W$, $E(v) = 0$. Logo, $v = E(v) = 0$ e, assim, $U \cap W = \{0\}$.

As duas propriedades demonstradas implicam que $V = U \oplus W$.

Problemas Complementares

Aplicações

5.45 Encontre o número de aplicações diferentes de (a) {1, 2} em {1, 2, 3}, (b) {1, 2, ..., r} em {1, 2, ..., s}.

5.46 Sejam $f : \mathbf{R} \to \mathbf{R}$ e $g : \mathbf{R} \to \mathbf{R}$ definidas por $f(x) = x^2 + 3x + 1$ e $g(x) = 2x - 3$. Encontre fórmulas que definam as aplicações compostas (a) $f \circ g$; (b) $g \circ f$; (c) $g \circ g$; (d) $f \circ f$.

5.47 Para cada aplicação $f : \mathbf{R} \to \mathbf{R}$, encontre uma fórmula para a inversa. (a) $f(x) = 3x - 7$, (b) $f(x) = x^3 + 2$.

5.48 Dada qualquer aplicação $f : A \to B$, mostre que $\mathbf{1}_B \circ f = f = f \circ \mathbf{1}_A$.

Transformações lineares

5.49 Mostre que são lineares as aplicações dadas.

(a) $F : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (x + 2y - 3z,\ 4x - 5y + 6z)$.

(b) $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (ax + by,\ cx + dy)$, com a, b, c, d escalares reais.

5.50 Mostre que não são lineares as aplicações dadas.

(a) $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (x^2, y^2)$.

(b) $F : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (x + 1,\ y + z)$.

(c) $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (xy, y)$.

(d) $F : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (|x|,\ y + z)$.

5.51 Encontre $F(a, b)$ se a transformação linear $F : \mathbf{R}^2 \to \mathbf{R}^2$ é definida por $F(1, 2) = (3, -1)$ e $F(0, 1) = (2, 1)$.

5.52 Encontre uma matriz 2×2 A que leve

(a) $(1, 3)^T$ e $(1, 4)^T$ em $(-2, 5)^T$ e $(3, -1)^T$, respectivamente.

(b) $(2, -4)^T$ e $(-1, 2)^T$ em $(1, 1)^T$ e $(1, 3)^T$, respectivamente.

5.53 Encontre uma matriz singular 2×2 B que leve $(1, 1)^T$ em $(1, 3)^T$.

5.54 Seja V o espaço vetorial de todas as matrizes reais quadradas de ordem n e seja M uma matriz não nula de V fixada. Mostre que as duas primeiras aplicações dadas são lineares, mas não a terceira. (a) $T(A) = MA$, (b) $T(A) = AM + MA$, (c) $T(A) = M + A$.

5.55 Dê um exemplo de uma aplicação não linear $F : \mathbf{R}^2 \to \mathbf{R}^2$ tal que $F^{-1}(0) = \{0\}$ mas F não é injetora.

5.56 Sejam $F : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (3x + 5y,\ 2x + 3y)$ e S o círculo unitário de $\mathbf{R}^2$. (S consiste nos pontos satisfazendo $x^2 + y^2 = 1$.) Encontre (a) a imagem $F(S)$, (b) a pré-imagem $F^{-1}(S)$.

5.57 Considere a transformação linear $G : \mathbf{R}^3 \to \mathbf{R}^3$ definida por $G(x, y, z) = (x + y + z,\ y - 2z,\ y - 3z)$ e a esfera unitária S_2 em $\mathbf{R}^3$, que consiste nos pontos satisfazendo $x^2 + y^2 + z^2 = 1$. Encontre (a) $G(S_2)$, (b) $G^{-1}(S_2)$.

5.58 Seja H o plano $x + 2y - 3z = 4$ de $\mathbf{R}^3$ e seja G a transformação linear do Problema 5.57. Encontre (a) $G(H)$, (b) $G^{-1}(H)$.

5.59 Seja W um subespaço de V. A aplicação inclusão $i : W \hookrightarrow V$ é definida por $i(w) = w$, para cada $w \in W$. Mostre que a inclusão é linear.

5.60 Seja $F : V \to U$ uma transformação linear. Mostre que $F(-v) = -F(v)$.

Núcleo e imagem de transformações lineares

5.61 Para cada uma das transformações lineares F dadas, encontre uma base e a dimensão do núcleo e da imagem de F.

(a) $F : \mathbf{R}^3 \to \mathbf{R}^3$ definida por $F(x, y, z) = (x + 2y - 3z,\ 2x + 5y - 4z,\ x + 4y + z)$,

(b) $F : \mathbf{R}^4 \to \mathbf{R}^3$ definida por $F(x, y, z, t) = (x + 2y + 3z + 2t,\ 2x + 4y + 7z + 5t,\ x + 2y + 6z + 5t)$.

5.62 Para cada uma das transformações lineares G dadas, encontre uma base e a dimensão do núcleo e da imagem de G.

(a) $G : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $G(x, y, z) = (x + y + z,\ 2x + 2y + 2z)$,

(b) $G : \mathbf{R}^3 \to \mathbf{R}^2$ definida por $G(x, y, z) = (x + y,\ y + z)$,

(c) $G : \mathbf{R}^5 \to \mathbf{R}^3$ definida por

$$G(x, y, z, s, t) = (x + 2y + 2z + s + t, \quad x + 2y + 3z + 2s - t, \quad 3x + 6y + 8z + 5s - t).$$

5.63 Cada uma das matrizes seguintes determina uma transformação linear de $\mathbf{R}^4$ em $\mathbf{R}^3$.

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{bmatrix}, \text{(b)} \quad B = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{bmatrix}.$$

Encontre uma base, bem como a dimensão, do núcleo e da imagem de cada transformação linear.

5.64 Encontre uma transformação linear $F : \mathbf{R}^3 \to \mathbf{R}^3$ cuja imagem seja gerada por (1, 2, 3) e (4, 5, 6).

5.65 Encontre uma transformação linear $G : \mathbf{R}^4 \to \mathbf{R}^3$ cujo núcleo seja gerado por (1, 2, 3, 4) e (0, 1, 1, 1).

5.66 Seja $V = \mathbf{P}_{10}(t)$ o espaço vetorial dos polinômios de grau ≤ 10. Considere a transformação linear $\mathbf{D}^4 : V \to V$, onde $\mathbf{D}^4$ denota a derivada quarta $d^4(f)/dt^4$. Encontre uma base e a dimensão (a) da imagem de $\mathbf{D}^4$; (b) do núcleo de $\mathbf{D}^4$.

5.67 Suponha que $F : V \to U$ seja linear. Mostre que (a) a imagem de qualquer subespaço de V é um subespaço de U; (b) a pré-imagem de qualquer subespaço de U é um subespaço de V.

5.68 Mostre que, se $F : V \to U$ é sobre, então $\dim U \leq \dim V$. Encontre todas as transformações lineares $F : \mathbf{R}^3 \to \mathbf{R}^4$ que são sobre.

5.69 Considere a transformação nula $0 : V \to U$ definida por $0(v) = 0$, para cada v de V. Encontre o núcleo e a imagem de 0.

Operações com transformações lineares

5.70 Sejam $F : \mathbf{R}^3 \to \mathbf{R}^2$ e $G : \mathbf{R}^3 \to \mathbf{R}^2$ definidas por $F(x, y, z) = (y,\ x + z)$ e $G(x, y, z) = (2z,\ x - y)$. Encontre fórmulas que definam as transformações $F + G$ e $3F - 2G$.

5.71 Seja $H : \mathbf{R}^2 \to \mathbf{R}^2$ definida por $H(x, y) = (y, 2x)$. Usando as transformações lineares F e G do Problema 5.70, encontre fórmulas que definam as transformações (a) $H \circ F$ e $H \circ G$, (b) $F \circ H$ e $G \circ H$, (c) $H \circ (F + G)$ e $H \circ F + H \circ G$.

5.72 Mostre que as transformações lineares F, G, H dadas são linearmente independentes.

(a) $F, G, H \in \text{Hom}(\mathbf{R}^2, \mathbf{R}^2)$ definidas por $F(x, y) = (x, 2y)$, $\quad G(x, y) = (y,\ x + y)$, $\quad H(x, y) = (0, x)$.

(b) $F, G, H \in \text{Hom}(\mathbf{R}^3, \mathbf{R})$ definidas por $F(x, y, z) = x + y + z$, $\quad G(x, y, z) = y + z$, $\quad H(x, y, z) = x - z$.

5.73 Dadas $F, G \in \text{Hom}(V, U)$, mostre que $\text{pos}(F + G) \leq \text{pos}(F) + \text{pos}(G)$. (Aqui, V tem dimensão finita.)

5.74 Sejam $F : V \to U$ e $G : U \to V$ lineares. Mostre que, se F e G são não singulares, então $G \circ F$ é não singular. Dê um exemplo em que $G \circ F$ é não singular, mas G não é. [*Sugestão*: considere $\dim V < \dim U$.]

5.75 Encontre a dimensão d de (a) $\text{Hom}(\mathbf{R}^2, \mathbf{R}^8)$, (b) $\text{Hom}(\mathbf{P}_4(t), \mathbf{R}^3)$, (c) $\text{Hom}(\mathbf{M}_{2,4}, \mathbf{P}_2(t))$.

5.76 Decida se a transformação linear dada é não singular. Se não for, encontre um vetor não nulo v cuja imagem seja 0; caso contrário, encontre uma fórmula para a transformação inversa.

(a) $F : \mathbf{R}^3 \to \mathbf{R}^3$ definida por $F(x, y, z) = (x + y + z, \quad 2x + 3y + 5z, \quad x + 3y + 7z)$,

(b) $G : \mathbf{R}^3 \to \mathbf{P}_2(t)$ definida por $G(x, y, z) = (x + y)t^2 + (x + 2y + 2z)t + y + z$,

(c) $H : \mathbf{R}^2 \to \mathbf{P}_2(t)$ definida por $H(x, y) = (x + 2y)t^2 + (x - y)t + x + y$.

5.77 Quando ocorre $\dim[\text{Hom}(V, U)] = \dim V$?

Álgebra dos operadores lineares

5.78 Sejam F e G os operadores lineares de $\mathbf{R}^2$ definidos por $F(x, y) = (x + y, 0)$ e $G(x, y) = (-y, x)$. Encontre as fórmulas que definem os operadores lineares (a) $F + G$, (b) $5F - 3G$, (c) FG, (d) GF, (e) F^2, (f) G^2.

5.79 Mostre que cada operador linear T de $\mathbf{R}^2$ dado é não singular e encontre uma fórmula para T^{-1}.
(a) $T(x, y) = (x + 2y, 2x + 3y)$, (b) $T(x, y) = (2x - 3y, \ 3x - 4y)$.

5.80 Mostre que cada um dos operadores lineares T de $\mathbf{R}^3$ dados é não singular e encontre uma fórmula para T^{-1}.
(a) $T(x, y, z) = (x - 3y - 2z, \ y - 4z, \ z)$; (b) $T(x, y, z) = (x + z, \ x - y, \ y)$.

5.81 Encontre a dimensão de $A(V)$, nos casos (a) $V = \mathbf{R}^7$, (b) $V = \mathbf{P}_5(t)$, (c) $V = \mathbf{M}_{3,4}$.

5.82 Quais dos inteiros dados pode ser a dimensão de uma álgebra $A(V)$ de operadores? 5, 9,12, 25, 28, 36, 45, 64, 88, 100.

5.83 Seja T o operador linear de $\mathbf{R}^2$ definido por $T(x, y) = (x + 2y, \ 3x + 4y)$. Encontre uma fórmula para $f(T)$, nos casos dados. (a) $f(t) = t^2 + 2t - 3$, (b) $f(t) = t^2 - 5t - 2$.

Problemas variados

5.84 Sejam $F : V \to U$ linear e k um escalar não nulo. Prove que as transformações F e kF têm o mesmo núcleo e a mesma imagem.

5.85 Sejam F e G operadores lineares de V. Suponha que F seja não singular e que a dimensão de V seja finita. Mostre que $\text{pos}(FG) = \text{pos}(GF) = \text{pos}(G)$.

5.86 Sejam V um espaço de dimensão finita e T um operador linear de V tal que $\text{pos}(T^2) = \text{pos}(T)$. Mostre que $\text{Nuc}\, T \cap \text{Im}\, T = \{0\}$.

5.87 Suponha que $V = U \oplus W$. Sejam E_1 e E_2 os operadores lineares de V definidos por $E_1(v) = u$, $E_2(v) = w$, com $v = u + w$, $u \in U$, $w \in W$. Mostre que (a) $E_1^2 = E_1$ e $E_2^2 = E_2$ (ou seja, E_1 e E_2 são projeções); (b) $E_1 + E_2 = I$, o operador identidade; (c) $E_1E_2 = 0$ e $E_2E_1 = 0$.

5.88 Sejam E_1 e E_2 operadores lineares de V quaisquer que satisfazem as condições (a), (b) e (c) do Problema 5.87. Prove que

$$V = \text{Im}\, E_1 \oplus \text{Im}\, E_2$$

5.89 Sejam v e w elementos de um espaço vetorial real V. O *segmento de reta* L de v a $v + w$ é definido como o conjunto de vetores $v + tw$, com $0 \le t \le 1$. (Ver Figura 5-6.)

(a) Mostre que o segmento de reta L entre os vetores v e u consiste nos pontos

$$\text{(i) } (1 - t)v + tu, \text{ com } 0 \le t \le 1, \quad \text{(ii) } t_1 v + t_2 u, \text{ com } t_1 + t_2 = 1, \ t_1 \ge 0, \ t_2 \ge 0.$$

(b) Seja $F : V \to U$ linear. Mostre que a imagem $F(L)$ de um segmento de reta L de V é um segmento de reta de U.

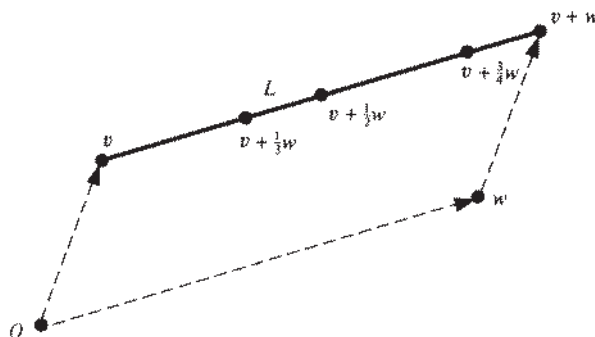



Figura 5-6

5.90 Seja $F : V \to U$ linear e W um subespaço de V. A *restrição* de F a W é a aplicação $F|W : W \to U$ definida por $F|W(v) = F(v)$, para cada v de W. Prove que

(a) $F|W$ é linear; (b) Nuc $(F|W) = (\text{Nuc } F) \cap W$; (c) Im $(F|W) = F(W)$.

5.91 Um subconjunto X de um espaço vetorial V é dito *convexo* se o segmento de reta entre dois pontos (vetores) $P, Q \in X$ quaisquer estiver contido em X. (a) Mostre que a interseção de conjuntos convexos é um conjunto convexo; (b) Suponha que $F : V \to U$ seja linear e X convexo. Mostre que $F(X)$ é convexo.

Respostas dos Problemas Complementares

5.45 (a) $3^2 = 9$, (b) s^r

5.46 (a) $(f \circ g)(x) = 4x^2 + 1$, (b) $(g \circ f)(x) = 2x^2 + 6x - 1$, (c) $(g \circ g)(x) = 4x - 9$,
(d) $(f \circ f)(x) = x^4 + 6x^3 + 14x^2 + 15x + 5$

5.47 (a) $f^{-1}(x) = \frac{1}{3}(x + 7)$, (b) $f^{-1}(x) = \sqrt[3]{x - 2}$

5.49 $F(x, y, z) = A(x, y, z)^T$, com (a) $A = \begin{bmatrix} 1 & 2 & -3 \\ 4 & -5 & 6 \end{bmatrix}$, (b) $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

5.50 (a) $u = (2, 2)$, $r = 3$; então $F(ru) = (36, 36)$, mas $rF(u) = (12, 12)$; (b) $F(0) \neq 0$;

(c) $u = (1, 2)$, $v = (3, 4)$; então $F(u + v) = (24, 6)$, mas $F(u) + F(v) = (14, 6)$;
(d) $u = (1, 2, 3)$, $r = -2$; então $F(ru) = (2, -10)$, mas $rF(u) = (-2, -10)$.

5.51 $F(a, b) = (-a + 2b, \ -3a + b)$

5.52 (a) $A = \begin{bmatrix} -17 & 5 \\ 23 & -6 \end{bmatrix}$; (b) Não existe. (2, –4) e (–1, 2) são linearmente dependentes, mas (1, 1) e (1, 3) não são.

5.53 $B = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$ [*Sugestão*: levar $(0, 1)^T$ em $(0, 0)^T$.]

5.55 $F(x, y) = (x^2, y^2)$

5.56 (a) $13x^2 - 42xy + 34y^2 = 1$, (b) $13x^2 + 42xy + 34y^2 = 1$

5.57 (a) $x^2 - 8xy + 26y^2 + 6xz - 38yz + 14z^2 = 1$, (b) $x^2 + 2xy + 3y^2 + 2xz - 8yz + 14z^2 = 1$

5.58 (a) $x - y + 2z = 4$, (b) $x + 6z = 4$

5.61 (a) $\dim(\text{Nuc}\, F) = 1$, $\{(7, -2, 1)\}$; $\dim(\text{Im}\, F) = 2$, $\{(1, 2, 1), \quad (0, 1, 2)\}$;
(b) $\dim(\text{Nuc}\, F) = 2$, $\{(-2, 1, 0, 0), \quad (1, 0, -1, 1)\}$; $\dim(\text{Im}\, F) = 2$, $\{(1, 2, 1), \quad (0, 1, 3)\}$

5.62 (a) $\dim(\text{Nuc}\, G) = 2$, $\{(1, 0, -1), \quad (1, -1, 0)\}$; $\dim(\text{Im}\, G) = 1$, $\{(1, 2)\}$;;
(b) $\dim(\text{Nuc}\, G) = 1$, $\{(1, -1, 1)\}$; $\text{Im}\, G = \mathbf{R}^2$, $\{(1, 0), \quad (0, 1)\}$;
(c) $\dim(\text{Nuc}\, G) = 3$, $\{(-2, 1, 0, 0, 0), \quad (1, 0, -1, 1, 0), \quad (-5, 0, 2, 0, 1)\}$; $\dim(\text{Im}\, G) = 2$, $\{(1, 1, 3), \quad (0, 1, 2)\}$

5.63 (a) $\dim(\text{Nuc}\, A) = 2$, $\{(4, -2, -5, 0), \quad (1, -3, 0, 5)\}$; $\dim(\text{Im}\, A) = 2$, $\{(1, 2, 1), \quad (0, 1, 1)\}$;
(b) $\dim(\text{Nuc}\, B) = 1$, $\{(-1, \frac{2}{3}, 1, 1)\}$; $\text{Im}\, B = \mathbf{R}^3$

5.64 $F(x, y, z) = (x + 4y, \; 2x + 5y, \; 3x + 6y)$

5.65 $F(x, y, z, t) = (x + y - z, \; 2x + y - t, 0)$

5.66 (a) $\{1, t, t^2, \ldots, t^6\}$, (b) $\{1, t, t^2, t^3\}$

5.68 Não há, porque $\dim \mathbf{R}^4 > \dim \mathbf{R}^3$.

5.69 $\text{Nuc}\; 0 = V$, $\text{Im}\; 0 = \{0\}$

5.70 $(F + G)(x, y, z) = (y + 2z, \; 2x - y + z)$, $(3F - 2G)(x, y, z) = (3y - 4z, \; x + 2y + 3z)$

5.71 (a) $(H \circ F)(x, y, z) = (x + z, \quad 2y)$, $(H \circ G)(x, y, z) = (x - y, \quad 4z)$; (b) não está definida;
(c) $(H \circ (F + G))(x, y, z) = (H \circ F + H \circ G)(x, y, z) = (2x - y + z, \quad 2y + 4z)$

5.74 $F(x, y) = (x, y, y)$, $G(x, y, z) = (x, y)$

5.75 (a) 16, (b) 15, (c) 24

5.76 (a) $v = (2, -3, 1)$; (b) $G^{-1}(at^2 + bt + c) = (b - 2c, \quad a - b + 2c, \quad -a + b - c)$;
(c) H é não singular, mas não invertível, porque $\mathbf{P}_2(t) > \dim \mathbf{R}^2$.

5.77 Se $\dim U = 1$, ou seja, se $U = K$.

5.78 (a) $(F + G)(x, y) = (x, x)$; (b) $(5F - 3G)(x, y) = (5x + 8y, \quad -3x)$; (c) $(FG)(x, y) = (x - y, \; 0)$;
(d) $(GF)(x, y) = (0, \; x + y)$; (e) $F^2(x, y) = (x + y, \quad 0)$ (observe que $F^2 = F$); (f) $G^2(x, y) = (-x, \quad -y)$.
[Observe que $G^2 + I = 0$, portanto, G é um zero de $f(t) = t^2 + 1$.]

5.79 (a) $T^{-1}(x, y) = (-3x + 2y, \; 2x - y)$, (b) $T^{-1}(x, y) = (-4x + 3y, \; -3x + 2y)$

5.80 (a) $T^{-1}(x, y, z) = (x + 3y + 14z, \; y - 4z, \; z)$, (b) $T^{-1}(x, y, z) = (y + z, \; y, \; x - y - z)$

5.81 (a) 49, (b) 36, (c) 144

5.82 Os quadrados: 9, 25, 36, 64 e 100

5.83 (a) $T(x, y) = (6x + 14y, \quad 21x + 27y)$; (b) $T(x, y) = (0, 0)$, ou seja, $f(T) = 0$

Capítulo 6

Transformações Lineares e Matrizes

6.1 INTRODUÇÃO

Considere uma base $S = \{u_1, u_2, ..., u_n\}$ de um espaço vetorial V sobre um corpo K. Dado um vetor $v \in V$ qualquer, digamos que

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

Então o vetor de coordenadas de v em relação à base S, que interpretamos como um vetor coluna (salvo menção explícita em contrário), é denotado e definido por

$$[v]_S = [a_1, a_2, \ldots, a_n]^T$$

Na Seção 4.11 vimos que a aplicação $v \mapsto [v]_S$, determinada pela base S, é um isomorfismo entre V e K^n.

Neste capítulo mostramos que também existe um isomorfismo, determinado pela base S, entre a álgebra $A(V)$ dos operadores lineares de V e a álgebra M das matrizes quadradas de ordem n sobre K. Assim, a cada operador linear $F: V \to V$ corresponde a uma matriz quadrada $[F]_S$ de ordem n determinada pela base S. Também veremos como nossa representação matricial varia quando escolhemos alguma outra base.

6.2 REPRESENTAÇÃO MATRICIAL DE UM OPERADOR LINEAR

Seja T um operador linear de um espaço vetorial V nele mesmo e suponha que $S = \{u_1, u_2, ..., u_n\}$ seja uma base de V. Agora $T(u_1), T(u_2), ..., T(u_n)$ são vetores de V, portanto, cada um é uma combinação linear dos vetores da base S, digamos

$$\begin{aligned} T(u_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ T(u_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ T(u_n) &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n \end{aligned}$$

Convém apresentar a definição seguinte.

DEFINIÇÃO A transposta da matriz de coeficientes $[a_{ij}]$, denotada por $m_S(T)$ ou $[T]_S$, é denominada *representação matricial* de T em relação à base S ou, simplesmente, matriz de T na base S. (O índice S pode ser omitido se soubermos qual é a base S.)

Usando a notação de vetores de coordenadas (colunas), a representação matricial de T pode ser escrita na forma

$$m_S(T) = [T]_S = \big[[T(u_1)]_S,\ [T(u_2)]_S,\ \ldots,\ [T(u_n)]_S\big]$$

Assim, as colunas de $m(T)$ são, respectivamente, os vetores de coordenadas de $T(u_1), T(u_2), ..., T(u_n)$.

Exemplo 6.1 Seja $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ o operador linear definido por $F(x, y) = (2x + 3y, 4x - 5y)$.

(a) Encontre a representação matricial de F em relação à base $S = \{u_1, u_2\} = \{(1, 2),\ (2, 5)\}$.

(1) Primeiro calculamos $F(u_1)$ e então escrevemos esse vetor como uma combinação linear dos vetores da base u_1 e u_2. (Por conveniência notacional, utilizamos vetores coluna.) Temos

$$F(u_1) = F\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}8\\-6\end{bmatrix} = x\begin{bmatrix}1\\2\end{bmatrix} + y\begin{bmatrix}2\\5\end{bmatrix} \qquad \text{e} \qquad \begin{aligned} x + 2y &= 8 \\ 2x + 5y &= -6 \end{aligned}$$

Resolvendo o sistema, obtemos $x = 52$ e $y = -22$. Logo, $F(u_1) = 52u_1 - 22u_2$.

(2) Em seguida calculamos $F(u_2)$ e então escrevemos esse vetor como uma combinação linear de u_1 e u_2. Temos

$$F(u_2) = F\left(\begin{bmatrix}2\\5\end{bmatrix}\right) = \begin{bmatrix}19\\-17\end{bmatrix} = x\begin{bmatrix}1\\2\end{bmatrix} + y\begin{bmatrix}2\\5\end{bmatrix} \qquad \text{e} \qquad \begin{aligned} x + 2y &= 19 \\ 2x + 5y &= -17 \end{aligned}$$

Resolvendo o sistema, obtemos $x = 129$ e $y = -55$. Logo, $F(u_2) = 129u_1 - 55u_2$.

Agora escrevemos as coordenadas de $F(u_1)$ e $F(u_2)$ como colunas para obter a matriz

$$[F]_S = \begin{bmatrix}52 & 129\\ -22 & -55\end{bmatrix}$$

(b) Encontre a representação matricial de F em relação à base canônica $E = \{e_1, e_2\} = \{(1, 0), (0, 1)\}$.

Encontramos $F(e_1)$ e escrevemos esse vetor como uma combinação linear dos vetores e_1 e e_2 da base canônica e, depois, encontramos $F(e_2)$ e escrevemos esse vetor como uma combinação linear de e_1 e e_2. Temos

$$\begin{aligned} F(e_1) &= F(1, 0) = (2, 2) \quad = 2e_1 + 4e_2 \\ F(e_2) &= F(0, 1) = (3, -5) = 3e_1 - 5e_2 \end{aligned} \qquad \text{e, portanto,} \qquad [F]_E = \begin{bmatrix}2 & 3\\ 4 & -5\end{bmatrix}$$

Observe que as coordenadas de $F(e_1)$ e $F(e_2)$ formam as colunas, e não as linhas, de $[F]_E$. Observe, também, que a aritmética é muito mais simples utilizando a base canônica de $\mathbf{R}^2$.

Exemplo 6.2 Seja V o espaço vetorial das funções, de base $S = \{\text{sen}\ t, \cos t, e^{3t}\}$, e seja $\mathbf{D}\colon V \to V$ o operador diferencial definido por $\mathbf{D}(f(t)) = d(f(t))/dt$. Calculemos a representação matricial de $\mathbf{D}$ na base S.

$$\begin{aligned} \mathbf{D}(\text{sen}\ t) &= \quad \cos t = \quad 0(\text{sen}\ t) + 1(\cos t) + 0(e^3 t) \\ \mathbf{D}(\cos\ t) &= -\text{sen}\ t = -1(\text{sen}\ t) + 0(\cos t) + 0(e^{3t}) \\ \mathbf{D}(e^{3t}) &= \quad 3e^{3t} = \quad 0(\text{sen}\ t) + 0(\cos t) + 3(e^{3t}) \end{aligned}$$

e, portanto,

$$[\mathbf{D}] = \begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 3\end{bmatrix}$$

Observe que as coordenadas de $\mathbf{D}(\text{sen}\ t)$, $\mathbf{D}(\cos t)$ e $\mathbf{D}(e^{3t})$ são as colunas, não as linhas, de $[\mathbf{D}]$.

Transformações matriciais e suas representações matriciais

Considere a matriz A dada a seguir, que pode ser vista como um operador linear de $\mathbf{R}^2$, e a base S de $\mathbf{R}^2$.

$$A = \begin{bmatrix}3 & -2\\ 4 & -5\end{bmatrix} \qquad \text{e} \qquad S = \{u_1, u_2\} = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix}\right\}$$

(Escrevemos os vetores como colunas, porque nossa aplicação é matricial.) Encontremos a representação matricial de A em relação à base S.

(1) Primeiro escrevemos $A(u_1)$ como uma combinação linear de u_1 e u_2. Temos

$$A(u_1) = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= -1 \\ 2x + 5y &= -6 \end{aligned}$$

Resolvendo o sistema, obtemos $x = 7$, $y = -4$. Assim, $A(u_1) = 7u_1 - 4u_2$.

(2) Em seguida, escrevemos $A(u_2)$ como uma combinação linear de u_1 e u_2. Temos

$$A(u_2) = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -4 \\ -7 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= -4 \\ 2x + 5y &= -7 \end{aligned}$$

Resolvendo o sistema, obtemos $x = -6$, $y = 1$. Assim, $A(u_2) = -6u_1 + u_2$. Escrevendo as coordenadas de $A(u_1)$ e $A(u_2)$ como colunas, obtemos a representação matricial de A, como segue.

$$[A]_S = \begin{bmatrix} 7 & -6 \\ -4 & 1 \end{bmatrix}$$

OBSERVAÇÃO Se quisermos encontrar a representação matricial de A em relação à base canônica $E = \{e_1, e_2\} = \{[1, 0]^T, [0, 1]^T\}$ de $\mathbf{R}^2$, calculamos

$$\begin{aligned} A(e_1) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 3e_1 + 4e_2 \\ A(e_2) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix} = -2e_1 - 5e_2 \end{aligned} \quad \text{e, portanto,} \quad [A]_E = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}$$

Observe que $[A]_E$ é a matriz original A. Esse resultado vale em geral.

> A representação matricial de qualquer matriz A $n \times n$ sobre K em relação à base canônica E de K^n é a própria matriz A, ou seja
> $$[A]_E = A$$

Algoritmo para determinar representações matriciais

Apresentamos a seguir um algoritmo para encontrar representações matriciais. O primeiro Passo 0 é opcional. Pode ser útil no Passo 1(b), que é repetido para cada vetor de base.

Algoritmo 6.1 Dados um operador linear T de um espaço vetorial V e uma base $S = \{u_1, u_2, ..., u_n\}$ de V, calculamos a representação matricial $[T]_S$.

Passo 0 Determine uma fórmula para as coordenadas de um vetor arbitrário v em relação à base S.

Passo 1 Repita, para cada vetor u_k de S,

(a) Calcule $T(u_k)$.

(b) Escreva $T(u_k)$ como uma combinação linear dos vetores $u_1, u_2, ..., u_n$ da base S.

Passo 2 Forme a matriz $[T]_S$, cujas colunas são os vetores de coordenadas do Passo 1(b).

Exemplo 6.3 Seja $F: \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (2x + 3y, 4x - 5y)$. Encontre a representação matricial $[F]_S$ de F em relação à base $S = \{u_1, u_2\} = \{(1, -2), (2, -5)\}$.

(Passo 0) Primeiro calculamos as coordenadas de $(a, b) \in \mathbf{R}^2$ em relação à base S. Temos

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 2 \\ -5 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 2y &= a \\ -2x - 5y &= b \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y &= a \\ -y &= 2a + b \end{aligned}$$

Resolvendo x e y em termos de a e b, obtemos $x = 5a + 2b$, $y = -2a - b$. Assim,

$$(a, b) = (5a + 2b)u_1 + (-2a - b)u_2$$

(Passo 1) Agora calculamos $F(u_1)$ e escrevemos esse vetor como uma combinação linear de u_1 e u_2 usando a fórmula de (a, b) e, depois, repetimos esse processo para $F(u_2)$. Temos

$$F(u_1) = F(1, -2) = \quad (-4, 14) = 8u_1 - 6u_2$$
$$F(u_2) = F(2, -5) = (-11, 33) = 11u_1 - 11u_2$$

(Passo 2) Finalmente, escrevemos as coordenadas de $F(u_1)$ e $F(u_2)$ como as colunas para obter a matriz procurada.

$$[F]_S = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix}$$

Propriedades da representação matricial

Nesta subseção mostramos as principais propriedades das representações matriciais de operadores lineares T de um espaço vetorial V. Enfatizamos que sempre estamos considerando dada uma base S de V específica dada.

Nosso primeiro teorema, demonstrado no Problema 6.9, nos diz que a "ação" de um operador linear T num vetor v é preservada por sua representação matricial.

Teorema 6.1 Sejam $T: V \to V$ um operador linear e S uma base (finita) de V. Então, para qualquer vetor v de V, $[T]_S[v]_S = [T(v)]_S$.

Exemplo 6.4 Considere o operador linear F de $\mathbf{R}^2$ e a base S do Exemplo 6.3, isto é,

$$F(x, y) = (2x + 3y, \quad 4x - 5y) \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1, -2), \quad (2, -5)\}$$

Seja

$$v = (5, -7), \quad \text{e, portanto,} \quad F(v) = (-11, 55)$$

Usando a fórmula do Exemplo 6.3, obtemos

$$[v] = [11, -3]^T \qquad \text{e} \qquad [F(v)] = [55, -33]^T$$

Verificamos o Teorema 6.1 para esse vetor v (onde $[F]$ foi obtida no Exemplo 6.3):

$$[F][v] = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix} \begin{bmatrix} 11 \\ -3 \end{bmatrix} = \begin{bmatrix} 55 \\ -33 \end{bmatrix} = [F(v)]$$

Dada uma particular base S de um espaço vetorial V, associamos uma matriz $[T]$ a cada operador linear T da álgebra $A(V)$ dos operadores lineares de V. O Teorema 6.1 nos diz que a "ação" de cada operador linear individual T é preservada por essa representação. Os próximos dois teoremas (demonstrados nos Problemas 6.10 e 6.11) nos dizem que também são preservadas as três operações básicas de $A(V)$ com esses operadores, a saber, (i) a adição, (ii) a multiplicação por escalar e (iii) a composição.

Teorema 6.2 Sejam V um espaço de dimensão n sobre K, S uma base de V e $\mathbf{M}$ a álgebra das matrizes $n \times n$ sobre K. Então a aplicação

$$m: A(V) \to \mathbf{M} \qquad \text{definida por} \qquad m(T) = [T]_S$$

é um isomorfismo de espaços vetoriais. Isto é, para quaisquer $F, G \in A(V)$ e $k \in K$, valem

(i) $m(F + G) = m(F) + m(G)$ ou $[F + G] = [F] + [G]$
(ii) $m(kF) = km(F)$ ou $[kF] = k[F]$
(iii) m é uma bijeção (injetora e sobre).

Teorema 6.3 Dados quaisquer operadores lineares $F, G \in A(V)$,

$$m(G \circ F) = m(G)m(F) \text{ or } [G \circ F] = [G][F]$$

(onde $G \circ F$ representa a composição das aplicações G e F).

6.3 MUDANÇA DE BASE

Seja V um espaço vetorial de dimensão n sobre um corpo K. Mostramos que, uma vez selecionada uma base S de V, cada vetor $v \in V$ pode ser representado por meio de uma ênupla $[v]_S$ em K^n e que cada operador linear T de $A(V)$ pode ser representado como uma matriz $n \times n$ sobre K. É natural formular a questão seguinte.

Como mudam essas representações se selecionarmos uma outra base?

Para responder essa pergunta, precisamos de mais uma definição.

DEFINIÇÃO Sejam $S = \{u_1, u_2, ..., u_n\}$ uma base do espaço vetorial V e $S' = \{v_1, v_2, \ldots, v_n\}$ uma outra base. (Para referência futura, dizemos que S é a base "antiga" e S' a base "nova".) Como S é uma base, cada vetor da base "nova" S' pode ser escrito, de maneira única, como uma combinação linear dos vetores de S, digamos,

$$\begin{aligned}
v_1 &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\
v_2 &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
v_n &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n
\end{aligned}$$

Seja P a transposta dessa matriz de coeficientes, ou seja, $P = [p_{ij}]$, onde $p_{ij} = a_{ji}$. Então P é denominada *matriz de mudança de base* (ou *matriz de transição*) da base "antiga" S para a base "nova" S'.

Destacamos as observações seguintes.

OBSERVAÇÃO 1 A matriz de mudança de base P pode ser vista como a matriz cujas colunas são, respectivamente, os vetores coluna de coordenadas dos "novos" vetores v_i da base em relação à base "antiga S, a saber,

$$P = \left[[v_1]_S, [v_2]_S, \ldots, [v_n]_S\right]$$

OBSERVAÇÃO 2 Analogamente, existe uma matriz de mudança de base Q da "nova" base S' para a base "antiga" S. Da mesma forma, Q pode ser vista como a matriz cujas colunas são, respectivamente, os vetores coluna de coordenadas dos "antigos" vetores u_i da base em relação à base "nova S', a saber,

$$Q = \left[[u_1]_{S'}, [u_2]_{S'}, \ldots, [u_n]_{S'}\right]$$

OBSERVAÇÃO 3 Como os vetores $v_1, v_2, ..., v_n$ da nova base S' são linearmente independentes, a matriz P é invertível (Problema 6.18). Analogamente, Q é invertível. De fato, a proposição seguinte é válida (demonstrada no Problema 6.18).

Proposição 6.4 Sejam P e Q as matrizes de mudança de base discutidas acima. Então $Q = P^{-1}$.

Sejam, agora, $\{u_1, u_2, ..., u_n\}$ uma base de um espaço vetorial V e $P = [p_{ij}]$ uma matriz não singular qualquer. Então os n vetores

$$v_i = p_{1i}u_i + p_{2i}u_2 + \cdots + p_{ni}u_n, \qquad i = 1, 2, \ldots, n$$

correspondentes às colunas de P, são linearmente independentes [Problema 6.21(a)]. Assim, determinam uma outra base S' de V. Além disso, P é a matriz de mudança de base de S para a nova base S'.

Exemplo 6.5 Considere as duas bases de $\mathbf{R}^2$ dadas a seguir.

$$S = \{u_1, u_2\} = \{(1,2),\ (3,5)\} \qquad \text{e} \qquad S' = \{v_1, v_2\} = \{(1,-1),\ (1,-2)\}$$

(a) Encontre a matriz P de mudança de base de S para a "nova" base S'.

Escrevemos cada um dos novos vetores de S' como uma combinação linear dos vetores u_1 e u_2 da base original S. obtendo

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{ou} \quad \begin{matrix} x + 3y = 1 \\ 2x + 5y = -1 \end{matrix} \quad \text{que fornece} \quad x = -8, \quad y = 3$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{ou} \quad \begin{matrix} x + 3y = 1 \\ 2x + 5y = -1 \end{matrix} \quad \text{que fornece} \quad x = -11, \quad y = 4$$

Assim,

$$\begin{aligned} v_1 &= \ \ -8u_1 + 3u_2 \\ v_2 &= -11u_1 + 4u_2 \end{aligned} \qquad \text{e, portanto,} \qquad P = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}.$$

Observe que as coordenadas de v_1 e v_2 são as colunas, e não as linhas, da matriz de mudança de base P.

(b) Encontre a matriz Q de mudança de base da "nova" base S' de volta à base "antiga" S.

Agora escrevemos cada um dos "antigos" vetores de base u_1 e u_2 de S como uma combinação linear dos "novos" vetores v_1 e v_2 da base S', obtendo

$$\begin{aligned} u_1 &= \ \ 4v_1 - 3v_2 \\ u_2 &= 11v_1 - 8v_2 \end{aligned} \qquad \text{e, portanto,} \qquad Q = \begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix}$$

Como era de se esperar pela Proposição 6.4, $Q = P^{-1}$. (De fato, poderíamos ter obtido Q simplesmente calculando P^{-1}.)

Exemplo 6.6 Considere as duas bases de $\mathbf{R}^3$ dadas a seguir.

$$E = \{e_1, e_2, e_3\} = \{(1,0,0), \quad (0,1,0), \quad (0,0,1)\}$$

e

$$S = \{u_1, u_2, u_3\} = \{(1,0,1), \quad (2,1,2), \quad (1,2,2)\}$$

(a) Encontre a matriz P de mudança da base E para a base S.

Como E é a base canônica, podemos escrever imediatamente cada elemento da base S como uma combinação linear dos vetores da base E. Mais precisamente,

$$\begin{aligned} u_1 &= (1,0,1) = \ \ e_1 + \qquad\ \ e_3 \\ u_2 &= (2,1,2) = 2e_1 + \ \ e_2 + 2e_3 \\ u_3 &= (1,2,2) = \ \ e_1 + 2e_2 + 2e_3 \end{aligned} \qquad \text{e, portanto,} \qquad P = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

Novamente, as coordenadas de u_1, u_2 e u_3 são as colunas de P. Observe que P simplesmente é a matriz cujas colunas são os vetores da base S. Isso só é válido porque a base original é a base canônica E.

(b) Encontre a matriz Q de mudança da base S para a base E.

A definição da matriz de mudança de base Q nos faz escrever cada um dos vetores da base canônica como uma combinação linear dos elementos da base S. Obtemos

$$\begin{aligned} e_1 &= (1,0,0) = -2u_1 + 2u_2 - u_3 \\ e_2 &= (0,1,0) = -2u_1 + \ \ u_2 \\ e_3 &= (0,0,1) = \ \ 3u_1 - 2u_2 + u_3 \end{aligned} \qquad \text{e, portanto,} \qquad Q = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

Enfatizamos que, para encontrar Q, precisamos resolver três sistemas de três equações lineares com três incógnitas, sendo um sistema 3×3 para cada um dos vetores e_1, e_2 e e_3.

Alternativamente, poderíamos ter calculado $Q = P^{-1}$ formando a matriz $M = [P, I]$ e reduzindo M à forma canônica por linhas.

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 & -2 & 3 \\ 0 & 1 & 0 & 2 & 1 & -2 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{bmatrix} = [I, P^{-1}]$$

Assim,

$$Q = P^{-1} = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

(Aqui utilizamos que Q é a inversa de P.)

O resultado do Exemplo 6.6(a) é geral. Como é muito utilizado, destacamos esse resultado formalmente.

Proposição 6.5 A matriz de mudança de base da base usual E de K^n para qualquer outra base S de K^n é a matriz P cujas colunas são, respectivamente, os vetores da base de S.

Aplicações da matriz de mudança de base

Inicialmente vejamos como uma mudança de base afeta as coordenadas de um vetor de um espaço vetorial V. O teorema seguinte está demonstrado no Problema 6.22.

Teorema 6.6 Seja P a matriz de mudança de base de uma base S para uma base S' de um espaço vetorial V. Então, dado um vetor $v \in V$ qualquer, temos

$$P[v]_{S'} = [v]_S \qquad \text{e, portanto,} \qquad P^{-1}[v]_S = [v]_{S'}$$

Em palavras, multiplicando as coordenadas de v na base original S por P^{-1}, obtemos as coordenadas de v na nova base S'.

OBSERVAÇÃO 1 Embora P seja dita a matriz de mudança de base da base antiga S para a nova base S', enfatizamos que é a matriz P^{-1} que transforma as coordenadas de v na base original S para as coordenadas de v na nova base S'.

OBSERVAÇÃO 2 Por causa do teorema, muitos livros consideram $Q = P^{-1}$, e não P, a matriz de mudança de base da base antiga S para a base nova S'. Alguns livros também se referem a Q como a *matriz de mudança de coordenadas*.

Vejamos, agora, a demonstração do teorema no caso particular em que dim $V = 3$. Seja P a matriz de mudança de base da base $S = \{u_1, u_2, u_3\}$ para a base $S' = \{v_1, v_2, v_3\}$;, digamos,

$$\begin{aligned} v_1 &= a_1u_1 + a_2u_2 + a_3a_3 \\ v_2 &= b_1u_1 + b_2u_2 + b_3u_3 \\ v_3 &= c_1u_1 + c_2u_2 + c_3u_3 \end{aligned} \qquad \text{e, portanto,} \qquad P = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

Seja, agora, $v \in V$ dado e digamos que $v = k_1v_1 + k_2v_2 + k_3v_3$. Então, substituindo as expressões de v_1, v_2, v_3, obtemos

$$\begin{aligned} v &= k_1(a_1u_1 + a_2u_2 + a_3u_3) + k_2(b_1u_1 + b_2u_2 + b_3u_3) + k_3(c_1u_1 + c_2u_2 + c_3u_3) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)u_1 + (a_2k_1 + b_2k_2 + c_2k_3)u_2 + (a_3k_1 + b_3k_2 + c_3k_3)u_3 \end{aligned}$$

Assim,

$$[v]_{S'} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} \quad \text{e} \quad [v]_S = \begin{bmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{bmatrix}$$

Dessa forma,

$$P[v]_{S'} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{bmatrix} = [v]_S$$

Finalmente, multiplicando a equação $[v]_S = P[v]_S$ por P^{-1}, obtemos

$$P^{-1}[v]_S = P^{-1}P[v]_{S'} = I[v]_{S'} = [v]_{S'}$$

O teorema seguinte (demonstrado no Problema 6.26) mostra como uma mudança de base afeta a representação matricial de um operador linear.

Teorema 6.7 Seja P a matriz de mudança de base de uma base S para uma base S' de um espaço vetorial V. Então, dado um operador linear T qualquer de V,

$$[T]_{S'} = P^{-1}[T]_S P$$

Ou seja, se A e B são as representações matriciais de T em relação às bases S e S', respectivamente, então

$$B = P^{-1}AP$$

Exemplo 6.7 Considere as duas bases de $\mathbf{R}^3$ dadas a seguir.

$$E = \{e_1, e_2, e_3\} = \{(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1)\}$$

e

$$S = \{u_1, u_2, u_3\} = \{(1, 0, 1), \quad (2, 1, 2), \quad (1, 2, 2)\}$$

A matriz de mudança de base P de E para S e sua inversa P^{-1} foram obtidas no Exemplo 6.6.

(a) Escreva $v = (1, 3, 5)$ como uma combinação linear de u_1, u_2, u_3 ou, equivalentemente, calcule $[v]_S$.

Uma maneira de fazer isso é resolver diretamente a equação vetorial $v = xu_1 + yu_2 + zu_3$, ou seja,

$$\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + z\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 1 \\ y + 2z &= 3 \\ x + 2y + 2z &= 5 \end{aligned}$$

A solução é $x = 7$, $y = –5$, $z = 4$, portanto, $v = 7u_1 - 5u_2 + 4u_3$.

Por outro lado, sabemos que $[v]_E = [1, 3, 5]^T$, porque E é a base canônica e já conhecemos P^{-1}. Portanto, pelo Teorema 6.6,

$$[v]_S = P^{-1}[v]_E = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ -5 \\ 4 \end{bmatrix}$$

Assim, novamente, $v = 7u_1 - 5u_2 + 4u_3$.

(b) Seja $A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix}$, que pode ser vista como um operador linear de $\mathbf{R}^3$. Encontre a matriz B que representa A em relação à base S.

A definição de representação matricial de A em relação à base S diz que devemos escrever cada um dos vetores $A(u_1)$, $A(u_2)$, $A(u_3)$ como uma combinação linear dos vetores u_1, u_2, u_3 de S. Obtemos

$$\begin{aligned} A(u_1) &= (-1,3,5) = 11u_1 - 5u_2 + 6u_3 \\ A(u_2) &= (1,2,9) \quad = 21u_1 - 14u_2 + 8u_3 \\ A(u_3) &= (3,-4,5) = 17u_1 - 8e_2 + 2u_3 \end{aligned} \qquad \text{e, portanto,} \qquad B = \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

Enfatizamos que, para encontrar B, precisamos resolver três sistemas de três equações lineares com três incógnitas, sendo um sistema 3×3 para cada um dos vetores $A(u_1)$, $A(u_2)$, $A(u_3)$.

Por outro lado, como conhecemos P e P^{-1}, podemos utilizar o Teorema 6.7, a saber,

$$B = P^{-1}AP = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

Como era de se esperar, obtemos o mesmo resultado.

6.4 SEMELHANÇA

Sejam A e B matrizes quadradas para as quais existe uma matriz invertível P tal que $B = P^{-1}AP$. Nesse caso, dizemos que B é *semelhante* a A ou então, obtida de A por uma *transformação de semelhança*. Mostramos (Problema 6.29) que a semelhança de matrizes é uma relação de equivalência.

Pelo Teorema 6.7 e essa última observação, obtemos o resultado a seguir.

Teorema 6.8 Duas matrizes representam o mesmo operador linear se, e só se, as matrizes são semelhantes.

Assim, todas as representações matriciais de um operador linear T formam uma classe de equivalência de matrizes semelhantes.

Um operador linear T é dito *diagonalizável* se existe uma base S de V na qual T é representado por uma matriz diagonal; nesse caso, dizemos que a base *S diagonaliza T*. O teorema precedente fornece o resultado seguinte.

Teorema 6.9 Seja A a representação matricial de um operador linear T. Então T é diagonalizável se, e só se, existe uma matriz invertível P tal que $P^{-1}AP$ é uma matriz diagonal.

Assim, T é diagonalizável se, e só se, sua representação matricial pode ser diagonalizada por uma transformação de semelhança.

Enfatizamos que nem todo operador é diagonalizável. Contudo, mostraremos (Capítulo 10) que todo operador linear pode ser representado por certas matrizes "padrão" denominadas formas *normais* ou *canônicas*. Para essa discussão, necessitamos de alguma teoria de corpos, polinômios e determinantes.

Funções e matrizes semelhantes

Seja f uma função de matrizes quadradas que associa o mesmo valor a matrizes semelhantes, isto é, $f(A) = f(B)$ sempre que A é semelhante a B. Então f induz, de uma maneira natural, uma função de operadores lineares T, também denotada por f, como segue. Definimos

$$f(T) = f([T]_S)$$

onde S é uma base qualquer. Pelo Teorema 6.8, essa função está bem definida.

Talvez o mais importante exemplo dessas funções seja o determinante (Capítulo 8). O traço (Seção 2.7) é um outro exemplo importante.

Exemplo 6.8 Considere o operador linear F e as bases E e S de $\mathbf{R}^2$ a seguir.

$$F(x,y) = (2x+3y,\ 4x-5y), \qquad E = \{(1,0),\ (0,1)\}, \qquad S = \{(1,2),\ (2,5)\}$$

Pelo Exemplo 6.1, as representações matriciais de F em relação às bases E e S são, respectivamente,

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} 52 & 129 \\ -22 & -55 \end{bmatrix}$$

Usando a matriz A, temos
(i) Determinante de $F = \det(A) = -10 - 12 = -22$; (ii) Traço de $F = \operatorname{tr}(A) = 2 - 5 = -3$.
Por outro lado, usando a matriz B, temos
(i) Determinante de $F = \det(B) = -2860 + 2838 = -22$; (ii) Traço de $F = \operatorname{tr}(B) = 52 - 55 = -3$.
Conforme esperado, ambas matrizes fornecem o mesmo resultado.

6.5 MATRIZES E TRANSFORMAÇÕES LINEARES ARBITRÁRIAS

Por último, consideramos o caso geral de transformações lineares de um espaço vetorial num outro. Sejam V e U espaços vetoriais sobre o mesmo corpo K e, digamos, que $\dim V = m$ e $\dim U = n$. Além disso, sejam

$$S = \{v_1, v_2, \ldots, v_m\} \qquad \text{e} \qquad S' = \{u_1, u_2, \ldots, u_n\}$$

respectivamente, bases arbitrárias fixadas de V e U.

Suponha que $F\colon V \to U$ seja uma transformação linear. Então os vetores $F(v_1), F(v_2), \ldots, F(v_m)$ pertencem a U e, portanto, cada um é uma combinação linear dos vetores da base S', digamos,

$$\begin{aligned} F(v_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ F(v_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F(v_m) &= a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n \end{aligned}$$

DEFINIÇÃO A transposta da matriz de coeficientes $[a_{ij}]$, denotada por $m_{S,S'}(F)$ ou $[F]_{S,S'}$, é denominada *representação matricial* de F em relação às bases S e S'. Utilizamos a notação mais simples $m(F)$ e $[F]$ quando as bases estiverem subentendidas.

O teorema seguinte é análogo ao Teorema 6.1 relativo a operadores lineares (Problema 6.67).

Teorema 6.10 Dado qualquer $v \in V$, temos $[F]_{S,S'}[v]_S = [F(v)]_{S'}$.

Assim, as coordenadas de $F(v)$ na base S' de U são obtidas pela multiplicação das coordenadas de v na base S de V por $[F]$.

Vimos que, dados quaisquer espaços vetoriais V e U, a coleção de todas as transformações lineares de V em U é um espaço vetorial denotado por $\operatorname{Hom}(V, U)$. O teorema seguinte é análogo ao Teorema 6.2 relativo a operadores lineares, em que, agora, $\mathbf{M} = \mathbf{M}_{m,n}$ denota o espaço vetorial de todas as matrizes $m \times n$ (Problema 6.67).

Teorema 6.11 A aplicação $m\colon \operatorname{Hom}(V,U) \to \mathbf{M}$ definida por $m(F) = [F]$ é um isomorfismo de espaços vetoriais, ou seja, dados quaisquer $F, G \in \operatorname{Hom}(V,U)$ e escalar k, temos

(i) $m(F+G) = m(F) + m(G)$ ou $[F+G] = [F] + [G]$
(ii) $m(kF) = km(F)$ ou $[kF] = k[F]$
(iii) m é uma bijeção (injetora e sobre).

Nosso próximo teorema é análogo ao Teorema 6.3 relativo a operadores lineares (Problema 6.67).

Teorema 6.12 Sejam S, S' e S'' bases dos espaços vetoriais V, U e W, respectivamente. Sejam $F: V \to U$ e $G \circ U \to W$ transformações lineares. Então

$$[G \circ F]_{S,S''} = [G]_{S',S''}[F]_{S,S'}$$

Ou seja, em relação a bases apropriadas, a representação matricial da composta de duas transformações lineares é a matriz produto das representações matriciais das transformações individuais.

Em seguida, veremos como a representação matricial de uma transformação linear $F: V \to U$ é afetada quando selecionamos bases novas.

Teorema 6.13 Seja P a matriz de mudança de base de uma base e para uma base e' de V e seja Q a matriz de mudança de base de uma base f para uma base f' de U. Então, dada qualquer transformação linear $F: V \to U$,

$$[F]_{e',f'} = Q^{-1}[F]_{e,f}P$$

Em outras palavras, se A é a representação matricial de uma transformação linear F relativa às bases e e f, e se B é a representação matricial de F nas bases e' e f', então

$$B = Q^{-1}AP$$

Nosso último teorema, demonstrado no Problema 6.36, mostra que toda transformação linear de um espaço vetorial V num outro espaço vetorial U pode ser representada por uma matriz muito simples. Observamos que esse teorema é análogo ao Teorema 3.18 relativo a matrizes $m \times n$.

Teorema 6.14 Seja $F: V \to U$ uma transformação linear, digamos, com pos(F) = r. Então existem bases de V e de U tais que a representação matricial de F é da forma

$$A = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

em que I_r é a matriz identidade de ordem r.

A matriz A desse teorema é denominada forma *normal* ou *canônica* da transformação linear F.

Problemas Resolvidos

Representação matricial de operadores lineares

6.1 Considere a transformação linear $F: \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (3x + 4y, \quad 2x - 5y)$ e as bases de $\mathbf{R}^2$ seguintes.

$$E = \{e_1, e_2\} = \{(1, 0), \ (0, 1)\} \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1, 2), \ (2, 3)\}$$

(a) Encontre a matriz A que representa F em relação à base E.

(b) Encontre a matriz B que representa F em relação à base S.

(a) Como E é a base canônica, as linhas de A são, simplesmente, os coeficientes dos componentes de $F(x, y)$, ou seja, usando $(a, b) = ae_1 + be_2$, temos

$$\begin{aligned} F(e_1) &= F(1, 0) = (3, 2) \quad = 3e_1 + 2e_2 \\ F(e_2) &= F(0, 1) = (4, -5) = 4e_1 - 5e_2 \end{aligned} \quad \text{e, portanto,} \quad A = \begin{bmatrix} 3 & 4 \\ 2 & -5 \end{bmatrix}$$

Observe que os coeficientes dos vetores da base estão escritos como colunas na representação matricial.

(b) Começamos calculando $F(u_1)$ e escrevendo esse vetor como uma combinação linear dos vetores u_1 e u_2 da base. Temos

$$F(u_1) = F(1,2) = (11, -8) = x(1,2) + y(2,3), \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= 11 \\ 2x + 3y &= -8 \end{aligned}$$

Resolvemos o sistema para obter $x = -49$, $y = 30$. Logo,

$$F(u_1) = -49u_1 + 30u_2$$

Em seguida, calculamos $F(u_2)$ e escrevemos esse vetor como uma combinação linear dos vetores u_1 e u_2. Temos

$$F(u_2) = F(2,3) = (18, -11) = x(1,2) + y(2,3), \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= 18 \\ 2x + 3y &= -11 \end{aligned}$$

Resolvemos em x e y e obtemos $x = -76$, $y = 47$. Logo,

$$F(u_2) = -76u_1 + 47u_2$$

Agora escrevemos os coeficientes de u_1 e u_2 como colunas, para obter $B = \begin{bmatrix} -49 & -76 \\ 30 & 47 \end{bmatrix}$

(b′) Alternativamente, podemos calcular primeiro as coordenadas de um vetor (a, b) arbitrário de $\mathbf{R}^2$ em relação à base S. Temos

$$(a, b) = x(1,2) + y(2,3) = (x + 2y,\ 2x + 3y), \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= a \\ 2x + 3y &= b \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = -3a + 2b$, $y = 2a - b$. Assim,

$$(a, b) = (-3a + 2b)u_1 + (2a - b)u_2$$

Em seguida, usamos a fórmula de (a, b) para encontrar as coordenadas de $F(u_1)$ e $F(u_2)$ em relação a S, como segue.

$$\begin{aligned} F(u_1) &= F(1,2) = (11, -8) = -49u_1 + 30u_2 \\ F(u_2) &= F(2,3) = (18, -11) = -76u_1 + 47u_2 \end{aligned} \quad \text{e, portanto,} \quad B = \begin{bmatrix} -49 & -76 \\ 30 & 47 \end{bmatrix}$$

6.2 Sejam G o operador linear de $\mathbf{R}^2$ e S a base dados a seguir.

$$G(x, y) = (2x - 7y,\ 4x + 3y) \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1,3),\ (2,5)\}$$

(a) Encontre a representação matricial $[G]_S$ de G em relação a S.

(b) Verifique $[G]_S[v]_S = [G(v)]_S$ para o vetor $v = (4, -3)$ de $\mathbf{R}^2$.

Começamos calculando as coordenadas de um vetor $v = (a, b)$ arbitrário de $\mathbf{R}^2$ em relação à base S, obtendo

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ 3 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= a \\ 3x + 5y &= b \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = -5a + 2b$, $y = 3a - b$. Assim,

$$(a, b) = (-5a + 2b)u_1 + (3a - b)u_2, \quad \text{e, portanto,} \quad [v] = [-5a + 2b, \quad 3a - b]^T$$

(a) Usando a fórmula de (a, b) e $G(x, y) = (2x - 7y,\ 4x + 3y)$, obtemos

$$\begin{aligned} G(u_1) &= G(1,3) = (-19, 13) = 121u_1 - 70u_2 \\ G(u_2) &= G(2,5) = (-31, 23) = 201u_1 - 116u_2 \end{aligned} \quad \text{e, portanto,} \quad [G]_S = \begin{bmatrix} 121 & 201 \\ -70 & -116 \end{bmatrix}$$

(Enfatizamos que os coeficientes de u_1 e u_2 aparecem como colunas, e não linhas, na representação matricial.)

(b) Usamos a fórmula $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$ para obter

$$\begin{aligned} v &= (4, -3) = -26u_1 + 15u_2 \\ G(v) &= G(4, -3) = (20, 7) = -131u_1 + 80u_2 \end{aligned}$$

Então $\qquad [v]_S = [-26, 15]^T \qquad \text{e} \qquad [G(v)]_S = [-131, 80]^T$

Assim,

$$[G]_S[v]_S = \begin{bmatrix} 121 & 201 \\ -70 & -116 \end{bmatrix}\begin{bmatrix} -26 \\ 15 \end{bmatrix} = \begin{bmatrix} -131 \\ 80 \end{bmatrix} = [G(v)]_S$$

(Isso é o esperado, segundo o Teorema 6.1.)

6.3 Sejam A a matriz quadrada de ordem 2 e S a base de $\mathbf{R}^2$ dadas a seguir.

$$A = \begin{bmatrix} 2 & 4 \\ 5 & 6 \end{bmatrix} \qquad \text{e} \qquad S = \{u_1, u_2\} = \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ -7 \end{bmatrix} \right\}$$

A matriz A define um operador linear de $\mathbf{R}^2$. Encontre a matriz B que representa a transformação A em relação à base S.

Começamos calculando as coordenadas de um vetor $(a, b)^T$ arbitrário em relação à base S, obtendo

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 3 \\ -7 \end{bmatrix} \qquad \text{ou} \qquad \begin{aligned} x + 3y &= a \\ -2x - 7y &= b \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = 7a + 3b$, $y = -2a - b$. Assim,

$$(a, b)^T = (7a + 3b)u_1 + (-2a - b)u_2$$

Em seguida, usamos a fórmula de $(a, b)^T$ para encontrar as coordenadas de Au_1 e Au_2 em relação a S, como segue.

$$Au_1 = \begin{bmatrix} 2 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -6 \\ -7 \end{bmatrix} = -63u_1 + 19u_2$$

$$Au_2 = \begin{bmatrix} 2 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ -7 \end{bmatrix} = \begin{bmatrix} -22 \\ -27 \end{bmatrix} = -235u_1 + 71u_2$$

Escrevendo as coordenadas como colunas, obtemos

$$B = \begin{bmatrix} -63 & -235 \\ 19 & 71 \end{bmatrix}$$

6.4 Encontre a representação matricial de cada um dos operadores lineares F de $\mathbf{R}^3$ em relação à base canônica $E = \{e_1, e_2, e_3\}$ de $\mathbf{R}^3$, ou seja, encontre $[F] = [F]_E$.

(a) F definido por $F(x, y, z) = (x + 2y - 3z,\ 4x - 5y - 6z,\ 7x + 8y + 9z)$.

(b) F definido pela matriz de ordem 3 $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 5 & 5 & 5 \end{bmatrix}$.

(c) F definido por $F(e_1) = (1, 3, 5), F(e_2) = (2, 4, 6), F(e_3) = (7, 7, 7)$. (O Teorema 5.2 afirma que uma transformação linear fica completamente definida por sua ação nos vetores de uma base.)

(a) Como E é a base canônica, simplesmente escrevemos os coeficientes dos componentes de $F(x, y, z)$ como colunas.

$$[F] = \begin{bmatrix} 1 & 2 & -3 \\ 4 & -5 & -6 \\ 7 & 8 & 9 \end{bmatrix}$$

(b) Como E é a base canônica, $[F] = A$, a própria matriz A.

(c) Aqui

$$\begin{aligned} F(e_1) &= (1, 3, 5) = e_1 + 3e_2 + 5e_3 \\ F(e_2) &= (2, 4, 6) = 2e_1 + 4e_2 + 6e_3 \\ F(e_3) &= (7, 7, 7) = 7e_1 + 7e_2 + 7e_3 \end{aligned} \qquad \text{e, portanto,} \quad [F] = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 4 & 7 \\ 5 & 6 & 7 \end{bmatrix}$$

Assim, as colunas de $[F]$ são as imagens dos vetores da base canônica.

6.5 Seja G o operador linear de $\mathbf{R}^3$ definido por $G(x, y, z) = (2y + z,\ x - 4y,\ 3x)$.

(a) Encontre a representação matricial de G em relação à base

$$S = \{w_1, w_2, w_3\} = \{(1, 1, 1), \quad (1, 1, 0), \quad (1, 0, 0)\}$$

(b) Verifique que $[G][v] = [G(v)]$, para cada vetor v de $\mathbf{R}^3$.

Começamos calculando as coordenadas de um vetor $(a, b, c) \in \mathbf{R}^3$ arbitrário em relação à base S. Escrevemos (a, b, c) como uma combinação linear de w_1, w_2, w_3 usando escalares x, y, z incógnitos, obtendo

$$(a, b, c) = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) = (x + y + z,\ \ x + y,\ \ x)$$

Igualando os componentes correspondentes, obtemos o sistema de equações

$$x + y + z = a, \qquad x + y = b, \qquad x = c$$

Resolvendo o sistema para x, y, z em termos de a, b, c obtemos $x = c$, $y = b - c$, $z = a - b$. Assim,

$$(a, b, c) = cw_1 + (b - c)w_2 + (a - b)w_3, \text{ ou, equivalentemente, } [(a, b, c)] = [c,\ \ b - c,\ \ a - b]^T$$

(a) Como $G(x, y, z) = (2y + z,\ \ x - 4y,\ \ 3x)$,

$$\begin{aligned} G(w_1) &= G(1, 1, 1) = (3, -3, 3) = 3w_1 - 6x_2 + 6x_3 \\ G(w_2) &= G(1, 1, 0) = (2, -3, 3) = 3w_1 - 6w_2 + 5w_3 \\ G(w_3) &= G(1, 0, 0) = (0, 1, 3) = 3w_1 - 2w_2 - w_3 \end{aligned}$$

Escrevemos as coordenadas $G(w_1)$, $G(w_2)$, $G(w_3)$ em colunas para obter

$$[G] = \begin{bmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{bmatrix}$$

(b) Escrevemos $G(v)$ como uma combinação linear de w_1, w_2, w_3, sendo $v = (a, b, c)$ um vetor arbitrário de $\mathbf{R}^3$.

$$G(v) = G(a, b, c) = (2b + c,\ \ a - 4b,\ \ 3a) = 3aw_1 + (-2a - 4b)w_2 + (-a + 6b + c)w_3$$

ou, equivalentemente,

$$[G(v)] = [3a,\ \ -2a - 4b,\ \ -a + 6b + c]^T$$

Assim,

$$[G][v] = \begin{bmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{bmatrix} \begin{bmatrix} c \\ b - c \\ a - b \end{bmatrix} = \begin{bmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{bmatrix} = [G(v)]$$

6.6 Sejam A a matriz quadrada de ordem 3 e S a base de $\mathbf{R}^3$ dadas a seguir.

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 3 & -1 & 0 \\ 1 & 4 & -2 \end{bmatrix} \qquad \text{e} \qquad S = \{u_1, u_2, u_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

A matriz A define um operador linear de $\mathbf{R}^3$. Encontre a matriz B que representa a transformação A em relação à base S. (Já sabemos que A representa a si mesma em relação à base canônica de $\mathbf{R}^3$.)

Começamos calculando as coordenadas de um vetor (a, b, c) de $\mathbf{R}^3$ arbitrário em relação à base S, obtendo

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + z \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \text{ou} \qquad \begin{aligned} x + z &= a \\ x + y + 2z &= b \\ x + y + 3z &= c \end{aligned}$$

Resolvemos x, y, z em termos de a, b, c para obter

$$x = a + b - c,\ y = -a + 2b - c,\ z = c - b$$

Assim, $\quad (a, b, c)^T = (a + b - c)u_1 + (-a + 2b - c)u_2 + (c - b)u_3$

Em seguida, usamos a fórmula de $(a, b, c)^T$ para encontrar as coordenadas de Au_1, Au_2, Au_3 em relação à base S, como segue.

$$\begin{aligned} A(u_1) &= A(1,1,1)^T = (0,2,3)^T &&= -u_1 + u_2 + u_3 \\ A(u_2) &= A(1,1,0)^T = (-1,-1,2)^T &&= -4u_1 - 3u_2 + 3u_3 \\ A(u_3) &= A(1,2,3)^T = (0,1,3)^T &&= -2u_1 - u_2 + 2u_3 \end{aligned}, \text{ portanto, } B = \begin{bmatrix} -1 & -4 & -2 \\ 1 & -3 & -1 \\ 1 & 3 & 2 \end{bmatrix}$$

6.7 Para cada uma das transformações lineares (operadores) L de $\mathbf{R}^2$ a seguir, encontre a matriz A que representa L (em relação à base canônica de $\mathbf{R}^2$).

(a) L é definida por $L(1, 0) = (2, 4)$ e $L(0, 1) = (5, 8)$.

(b) L é a rotação de $\mathbf{R}^2$ no sentido anti-horário por $90°$.

(c) L é a reflexão de $\mathbf{R}^2$ em torno da reta $y = -x$.

(a) Como $\{(1, 0), (0, 1)\}$ é a base canônica de $\mathbf{R}^2$, escrevemos as imagens por L como colunas para obter

$$A = \begin{bmatrix} 2 & 5 \\ 4 & 8 \end{bmatrix}$$

(b) Pela rotação L, temos $L(1, 0) = (0, 1)$ e $L(0, 1) = (-1, 0)$. Assim,

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

(c) Pela reflexão L, temos $L(1, 0) = (0, -1)$ e $L(0, 1) = (-1, 0)$. Assim,

$$A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

6.8 O conjunto $S = \{e^{3t}, te^{3t}, t^2e^{3t}\}$ é uma base do espaço vetorial V de funções $f: \mathbf{R} \to \mathbf{R}$. Seja $\mathbf{D}$ o operador derivada de V, ou seja, $\mathbf{D}(f) = df/dt$. Encontre a representação matricial de $\mathbf{D}$ em relação à base S.

Calculamos a imagem de cada função da base, como segue.

$$\begin{aligned} \mathbf{D}(e^{3t}) &= 3e^{3t} &&= 3(e^{3t}) + 0(te^{3t}) + 0(t^2e^{3t}) \\ \mathbf{D}(te^{3t}) &= e^{3t} + 3te^{3t} &&= 1(e^{3t}) + 3(te^{3t}) + 0(t^2e^{3t}) \\ \mathbf{D}(t^2e^{3t}) &= 2te^{3t} + 3t^2e^{3t} &&= 0(e^{3t}) + 2(te^{3t}) + 3(t^2e^{3t}) \end{aligned} \quad \text{e, portanto,} \quad [\mathbf{D}] = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

6.9 Demonstre o Teorema 6.1. Sejam $T: V \to V$ um operador linear e S uma base (finita) de V. Então, para qualquer vetor v de V, $[T]_S[v]_S = [T(v)]_S$.

Seja $S = \{u_1, u_2, ..., u_n\}$ e suponha que, para $i = 1, ..., n$,

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j$$

Então $[T]_S$ é a matriz quadrada de ordem n cuja j-ésima linha é

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Agora, vamos supor que

$$v = k_1u_1 + k_2u_2 + \cdots + k_nu_n = \sum_{i=1}^{n} k_iu_i$$

Escrevendo um vetor coluna como a transposta de um vetor linha, temos

$$[v]_S = [k_1, k_2, \ldots, k_n]^T \tag{2}$$

Além disso, usando a linearidade de T,

$$T(v) = T\left(\sum_{i=1}^{n} k_i u_i\right) = \sum_{i=1}^{n} k_i T(u_i) = \sum_{i=1}^{n} k_i \left(\sum_{j=1}^{n} a_{ij} u_j\right)$$
$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij} k_i\right) u_j = \sum_{j=1}^{n} (a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n) u_j$$

Assim, $[T(v)]_S$ é o vetor coluna cuja j-ésima entrada é

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \tag{3}$$

Por outro lado, a j-ésima entrada de $[T]_S[v]_S$ é obtida multiplicando a j-ésima linha de $[T]_S$ por $[v]_S$, ou seja, (1) por (2). Mas o produto de (1) por (2) é (3). Logo, $[T]_S[v]_S$ e $[T(v)]_S$ têm as mesmas entradas. Assim, $[T]_S[v]_S = [T(v)]_S$.

6.10 Demonstre o Teorema 6.2. Sejam $S = \{u_1, u_2, ..., u_n\}$ uma base de V sobre K e $\mathbf{M}$ a álgebra das matrizes quadradas de ordem n sobre K. Então a aplicação $m: A(V) \to \mathbf{M}$ definida por $m(T) = [T]_S$ é um isomorfismo de espaços vetoriais. Isto é, para quaisquer $F, G \in A(V)$ e $k \in K$, valem

(i) $[F + G] = [F] + [G]$, (ii) $[kF] = k[F]$, (iii) m é injetora e sobre.

(i) Suponha que, para $i = 1, ..., n$,

$$F(u_i) = \sum_{j=1}^{n} a_{ij} u_j \qquad \text{e} \qquad G(u_i) = \sum_{j=1}^{n} b_{ij} u_j$$

Considere as matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$. Então $[F] = A^T$ e $[G] = B^T$. Para $i = 1, ..., n$, temos

$$(F + G)(u_i) = F(u_i) + G(u_i) = \sum_{j=1}^{n} (a_{ij} + b_{ij}) u_j$$

Como $A + B$ é a matriz $[a_{ij} + b_{ij}]$, temos

$$[F + G] = (A + B)^T = A^T + B^T = [F] + [G]$$

(ii) Para $i = 1, ..., n$,

$$(kF)(u_i) = kF(u_i) = k\sum_{j=1}^{n} a_{ij} u_j = \sum_{j=1}^{n} (ka_{ij}) u_j$$

Como kA é a matriz $[ka_{ij}]$, temos

$$[kF] = (kA)^T = kA^T = k[F]$$

(iii) Finalmente, m é injetora, porque uma transformação linear fica completamente determinada por seus valores numa base. Também m é sobre, porque a matriz $A = [a_{ij}]$ de $\mathbf{M}$ é a imagem do operador linear

$$F(u_i) = \sum_{j=1}^{n} a_{ij} u_j, \qquad i = 1, \ldots, n$$

Assim, o teorema está demonstrado.

6.11 Demonstre o Teorema 6.3. Dados quaisquer operadores lineares $G, F \in A(V)$, vale $[G \circ F] = [G][F]$.

Usando a notação do Problema 6.10, temos

$$(G \circ F)(u_i) = G(F(u_i)) = G\left(\sum_{j=1}^{n} a_{ij} u_j\right) = \sum_{j=1}^{n} a_{ij} G(u_j)$$
$$= \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk} u_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) u_k$$

Lembre que AB é a matriz $AB = [c_{ik}]$, com $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$. Por isso,

$$[G \circ F] = (AB)^T = B^T A^T = [G][F]$$

Isso prova o teorema.

6.12 Seja A a representação matricial de um operador linear T. Prove que, dado qualquer polinômio $f(t)$, temos que $f(A)$ é a representação matricial de $f(T)$. [Assim, $f(T) = 0$ se, e só se, $f(A) = 0$.]

Seja ϕ a aplicação que leva cada operador T em sua representação matricial A. Queremos provar que $\phi(f(T)) = f(A)$. Suponha que $f(t) = a_n t^n + \cdots + a_1 t + a_0$. A demonstração é por indução em n, o grau de $f(t)$.

Suponha que $n = 0$. Lembre que $\phi(I') = I$, onde I' é a transformação identidade e I é a matriz identidade. Assim,

$$\phi(f(T)) = \phi(a_0 I') = a_0 \phi(I') = a_0 I = f(A)$$

e, portanto, o teorema vale para $n = 0$.

Agora supomos que o teorema valha para polinômios de grau menor do que n. Então, como ϕ é um isomorfismo de álgebras,

$$\begin{aligned}\phi(f(T)) &= \phi(a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') \\ &= a_n \phi(T)\phi(T^{n-1}) + \phi(a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') \\ &= a_n A A^{n-1} + (a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I) = f(A)\end{aligned}$$

e terminamos a demonstração do teorema.

Mudança de base

Nesta seção, o vetor de coordenadas $[v]_S$ sempre denota um vetor coluna, ou seja,

$$[v]_S = [a_1, a_2, \ldots, a_n]^T$$

6.13 Considere as bases de $\mathbf{R}^2$ a seguir.

$$E = \{e_1, e_2\} = \{(1,0),\ (0,1)\} \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1,3),\ (1,4)\}$$

(a) Encontre a matriz de mudança de base P da base canônica E para S.

(b) Encontre a matriz de mudança de base Q de S de volta para E.

(c) Encontre o vetor de coordenadas $[v]$ de $v = (5, -3)$ em relação a S.

(a) Como E é a base canônica, simplesmente escrevemos os vetores da base S como colunas: $P = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$

(b) **Método 1** Usamos a definição de matriz de mudança de base, ou seja, escrevemos cada vetor de E como uma combinação linear dos vetores de S. Para isso, começamos calculando as coordenadas de um vetor $v = (a, b)$ arbitrário em relação a S. Temos

$$(a,b) = x(1,3) + y(1,4) = (x + y, 3x + 4y) \qquad \text{ou} \qquad \begin{aligned} x + \ \ y &= a \\ 3x + 4y &= b \end{aligned}$$

Resolvemos em x e y para obter $x = 4a - b$, $y = -3a + b$. Assim,

$$v = (4a - b)u_1 + (-3a + b)u_2 \qquad \text{e} \qquad [v]_S = [(a,b)]_S = [4a - b,\ -3a + b]^T$$

Usando essa fórmula de $[v]_S$ e escrevendo as coordenadas dos e_i como colunas, obtemos

$$\begin{aligned} e_1 &= (1,0) = \ 4u_1 - 3u_2 \\ e_2 &= (0,1) = -u_1 + \ u_2 \end{aligned} \qquad \text{e} \qquad Q = \begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix}$$

Método 2 Como $Q = P^{-1}$, encontramos P^{-1}, digamos, usando a fórmula da inversa de matrizes 2×2. Assim,

$$P^{-1} = \begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix}$$

(c) **Método 1** Escrevemos v como uma combinação linear dos vetores de S, digamos, usando a fórmula de $v = (a, b)$. Temos $v = (5, -3) = 23u_1 - 18u_2$ e, portanto, $[v]_S = [23, -18]^T$.

Método 2 Pelo Teorema 6.6, podemos usar $[v]_S = P^{-1}[v]_E$ e $[v]_E = [5, -3]^T$ para obter

$$[v]_S = P^{-1}[v]_E = \begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \end{bmatrix} = \begin{bmatrix} 23 \\ -18 \end{bmatrix}$$

6.14 Os vetores $u_1 = (1, 2, 0)$, $u_2 = (1, 3, 2)$, $u_3 = (0, 1, 3)$ formam uma base S de $\mathbf{R}^3$. Encontre

(a) a matriz de mudança de base P da base canônica $E = \{e_1, e_2, e_3\}$ para S.

(b) Encontre a matriz de mudança de base Q de S de volta para E.

(a) Como E é a base canônica, simplesmente escrevemos os vetores da base S como colunas: $P = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 2 & 3 \end{bmatrix}$.

(b) **Método 1** Escrevemos cada vetor da base E como uma combinação linear dos vetores da base S, calculando as coordenadas de um vetor $v = (a, b, c)$ arbitrário em relação à base S. Temos

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + z \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + y &= a \\ 2x + 3y + z &= b \\ 2y + 3z &= c \end{aligned}$$

Resolvemos em x, y, z para obter $x = 7a - 3b + c$, $y = -6a + 3b - c$, $z = 4a - 2b + c$. Assim,

$$v = (a, b, c) = (7a - 3b + c)u_1 + (-6a + 3b - c)u_2 + (4a - 2b + c)u_3$$

ou $$[v]_S = [(a, b, c)]_S = [7a - 3b + c, \ -6a + 3b - c, \ 4a - 2b + c]^T$$

Usando essa fórmula de $[v]_S$ e escrevendo as coordenadas dos e_i como colunas, obtemos

$$\begin{aligned} e_1 &= (1, 0, 0) = 7u_1 - 6u_2 + 4u_3 \\ e_2 &= (0, 1, 0) = -3u_1 + 3u_2 - 2u_3 \\ e_3 &= (0, 0, 1) = u_1 - u_2 + u_3 \end{aligned} \quad \text{e} \quad Q = \begin{bmatrix} 7 & -3 & 1 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{bmatrix}$$

Método 2 Encontramos P^{-1} reduzindo $M = [P, I]$ à forma $[I, P^{-1}]$, como segue.

$$M = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7 & -3 & 1 \\ 0 & 1 & 0 & -6 & 3 & -1 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right] = [I, P^{-1}]$$

Assim, $Q = P^{-1} = \begin{bmatrix} 7 & -3 & 1 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{bmatrix}$.

6.15 Suponha que os eixos x e y do plano $\mathbf{R}^2$ tenham sido girados por 45° no sentido anti-horário, de modo que os novos eixos x' e y' estejam, agora, na posição das retas $y = x$ e $y = -x$, respectivamente.

(a) Encontre a matriz P de mudança de base.

(b) Encontre as coordenadas do ponto $A(5, 6)$ com a rotação dada.

(a) Os vetores unitários na direção dos novos eixos x' e y' são

$$u_1 = (\tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2}) \qquad \text{e} \qquad u_2 = (-\tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2})$$

(Os vetores unitários na direção dos eixos x e y originais são os vetores da base canônica de $\mathbf{R}^2$.) Assim, escrevemos as coordenadas de u_1 e u_2 como colunas para obter

$$P = \begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

(b) Multiplicamos as coordenadas do ponto dado por P^{-1}, como segue.

$$\begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{11}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix}$$

(Como P é ortogonal, a inversa P^{-1} é, simplesmente, a transposta de P.)

6.16 Os vetores $u_1 = (1, 1, 0)$, $u_2 = (0, 1, 1)$, $u_3 = (1, 2, 2)$ formam uma base S de $\mathbf{R}^3$. Encontre as coordenadas de um vetor arbitrário $v = (a, b, c)$ em relação à base S.

Método 1 Escrevemos v como uma combinação linear de u_1, u_2, u_3 usando incógnitas x, y, z. Temos

$$(a, b, c) = x(1, 1, 0) + y(0, 1, 1) + z(1, 2, 2) = (x + z,\ \ x + y + 2z,\ \ y + 2z)$$

que fornece o sistema

$$\begin{aligned} x + \quad\ z &= a \\ x + y + 2z &= b \\ y + 2z &= c \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + \quad\ z &= a \\ y + \ z &= -a + b \\ y + 2z &= c \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + \quad z &= a \\ y + z &= -a + b \\ z &= a - b + c \end{aligned}$$

Substituindo para trás, obtemos $x = b - c$, $y = -2a + 2b - c$, $z = a - b + c$. Assim,

$$[v]_S = [b - c,\ \ -2a + 2b - c,\ \ a - b + c]^T$$

Método 2 Encontramos P^{-1} reduzindo $M = [P, I]$ à forma $[I, P^{-1}]$, em que P é a matriz de mudança de base da base canônica E para S ou, em outras palavras, a matriz cujas colunas são os vetores da base S.

Temos

$$M = \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -2 & 2 & -1 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right] = [I, P^{-1}]$$

Assim, $$P^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix} \quad \text{e} \quad [v]_S = P^{-1}[v]_E = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b - c \\ -2a + 2b - c \\ a - b + c \end{bmatrix}$$

6.17 Considere as bases de $\mathbf{R}^2$ a seguir.

$$S = \{u_1, u_2\} = \{(1, -2),\ (3, -4)\} \qquad \text{e} \qquad S' = \{v_1, v_2\} = \{(1, 3),\ (3, 8)\}$$

(a) Encontre as coordenadas de $v = (a, b)$ em relação à base S.

(b) Encontre a matriz de mudança de base P de S para S'.

(c) Encontre as coordenadas de $v = (a, b)$ em relação à base S'.

(d) Encontre a matriz de mudança de base Q de S' de volta para S.

(e) Verifique que $Q = P^{-1}$.

(f) Mostre que $P^{-1}[v]_S = [v]_{S'}$, para qualquer vetor $v = (a, b)$ de $\mathbf{R}^2$. (Ver Teorema 6.6.)

(a) Seja $v = xu_1 + yu_2$, com incógnitas x e y, ou seja,

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 3 \\ -4 \end{bmatrix} \qquad \text{ou} \qquad \begin{aligned} x + 3y &= a \\ -2x - 4y &= b \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 3y &= a \\ 2y &= 2a + b \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = -2a - \frac{3}{2}b$ e $y = a + \frac{1}{2}b$. Assim,

$$(a, b) = (-2a - \tfrac{3}{2})u_1 + (a + \tfrac{1}{2}b)u_2 \qquad \text{ou} \qquad [(a, b)]_S = [-2a - \tfrac{3}{2}b,\ \ a + \tfrac{1}{2}b]^T$$

(b) Usamos a parte (a) para escrever cada vetor v_1 e v_2 da base S' como uma combinação linear dos vetores u_1 e u_2 da base S, ou seja,

$$v_1 = (1, 3) = (-2 - \tfrac{9}{2})u_1 + (1 + \tfrac{3}{2})u_2 = -\tfrac{13}{2}u_1 + \tfrac{5}{2}u_2$$
$$v_2 = (3, 8) = (-6 - 12)u_1 + (3 + 4)u_2 = -18u_1 + 7u_2$$

Então P é a matriz cujas colunas são as coordenadas de v_1 e v_2 em relação à base S, ou seja,

$$P = \begin{bmatrix} -\frac{13}{2} & -18 \\ \frac{5}{2} & 7 \end{bmatrix}$$

(c) Seja $v = xv_1 + yv_2$, com incógnitas escalares x e y, ou seja,

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ 3 \end{bmatrix} + y\begin{bmatrix} 3 \\ 8 \end{bmatrix} \quad \text{ou} \quad \begin{aligned} x + 3y &= a \\ 3x + 8y &= b \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 3y &= a \\ -y &= b - 3a \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = -8a + 3b$ e $y = 3a - b$. Assim,

$$(a, b) = (-8a + 3b)v_1 + (3a - b)v_2 \qquad \text{ou} \qquad [(a, b)]_{S'} = [-8a + 3b, \quad 3a - b]^T$$

(d) Usamos a parte (c) para escrever cada vetor u_1 e u_2 da base S como uma combinação linear dos vetores v_1 e v_2 da base S', ou seja,

$$\begin{aligned} u_1 &= (1, -2) = (-8 - 6)v_1 + (3 + 2)v_2 = -14v_1 + 5v_2 \\ u_2 &= (3, -4) = (-24 - 12)v_1 + (9 + 4)v_2 = -36v_1 + 13v_2 \end{aligned}$$

Escrevemos as coordenadas de u_1 e u_2 em relação a S' como colunas para obter $Q = \begin{bmatrix} -14 & -36 \\ 5 & 13 \end{bmatrix}$.

(e) $QP = \begin{bmatrix} -14 & -36 \\ 5 & 13 \end{bmatrix}\begin{bmatrix} -\frac{13}{2} & -18 \\ \frac{5}{2} & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

(f) Usando as partes (a), (c) e (d), obtemos

$$P^{-1}[v]_S = Q[v]_S = \begin{bmatrix} -14 & -36 \\ 5 & 13 \end{bmatrix}\begin{bmatrix} -2a - \frac{3}{2}b \\ a + \frac{1}{2}b \end{bmatrix} = \begin{bmatrix} -8a + 3b \\ 3a - b \end{bmatrix} = [v]_{S'}$$

6.18 Sejam P a matriz de mudança de base de uma base $\{u_i\}$ para uma base $\{w_i\}$ e Q a matriz de mudança de base da base $\{w_i\}$ de volta para a base $\{u_i\}$. Prove que P é invertível e que $Q = P^{-1}$.

Suponha, para $i = 1, 2, ..., n$, que

$$w_i = a_{i1}u_1 + a_{i2}u_2 + \ldots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j \tag{1}$$

e que, para $j = 1, 2, ..., n$,

$$u_j = b_{j1}w_1 + b_{j2}w_2 + \cdots + b_{jn}w_n = \sum_{k=1}^{n} b_{jk}w_k \tag{2}$$

Sejam $A = [a_{ij}]$ e $B = [b_{jk}]$. Então $P = A^T$ e $Q = B^T$. Substituindo (2) em (1), obtemos

$$w_i = \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk}w_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)w_k$$

Como $\{w_i\}$ é uma base, $\sum a_{ij}b_{jk} = \delta_{ik}$, onde δ_{ik} é o delta de Kronecker, ou seja, $\delta_{ik} = 1$ se $i = k$ e $\delta_{ik} = 0$ se $i \neq k$. Suponha que $AB = [c_{ik}]$. Então $c_{ik} = \delta_{ik}$. Por isso, $AB = I$ e, portanto,

$$QP = B^T A^T = (AB)^T = I^T = I$$

Assim, $Q = P^{-1}$.

6.19 Considere uma sequência finita de vetores $S = \{u_1, u_2, ..., u_n\}$. Seja S' a sequência de vetores obtida de S por alguma das "operações elementares" a seguir.

(1) Trocar dois vetores de posição.

(2) Multiplicar um vetor por um escalar não nulo.

(3) Somar um múltiplo de um vetor a um outro vetor.

Mostre que S e S' geram o mesmo subespaço W. Mostre também que S' é linearmente independente se, e só se, S é linearmente independente.

Observe que, com cada uma das operações, os vetores de S' são combinações lineares dos vetores de S. Como cada operação tem uma inversa do mesmo tipo, cada vetor de S é uma combinação linear dos vetores de S'. Assim, S e S' geram o mesmo subespaço W. Além disso, S' é linearmente independente se, e só se, $\dim W = n$, e isso é válido se, e só se, S é linearmente independente.

6.20 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes $m \times n$ sobre um corpo K equivalentes por linhas e sejam $v_1, v_2, ..., v_n$ vetores quaisquer de um espaço vetorial V sobre K. Dado $i = 1, 2, ..., m$, sejam u_i e w_i os vetores definidos por

$$u_i = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n \qquad \text{e} \qquad w_i = b_{i1}v_1 + b_{i2}v_2 + \cdots + b_{in}v_n$$

Mostre que $\{u_i\}$ e $\{w_i\}$ geram o mesmo subespaço de V.

Uma "operação elementar" do Problema 6.19 aplicada a $\{u_i\}$ equivale a uma operação elementar com as linhas de A. Como A e B são equivalentes por linhas, B pode ser obtida a partir de A com uma sequência de operações elementares com as linhas. Logo, $\{w_i\}$ pode ser obtido a partir de $\{u_i\}$ pela correspondente sequência de "operações elementares" e, por isso, $\{u_i\}$ e $\{w_i\}$ geram o mesmo subespaço.

6.21 Sejam $u_1, u_2, ..., u_n$ vetores de um espaço vetorial V sobre K e $P = [p_{ij}]$ uma matriz quadrada de ordem n sobre K. Dado $i = 1, 2, ..., n$, seja $v_i = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n$.

(a) Suponha que P seja invertível. Mostre que $\{u_i\}$ e $\{v_i\}$ geram o mesmo subespaço de V. Logo, $\{u_i\}$ é linearmente independente se, e só se, $\{v_i\}$ é linearmente independente.

(b) Suponha que P seja singular (não invertível). Mostre que $\{v_i\}$ é linearmente dependente.

(c) Suponha que $\{v_i\}$ seja linearmente independente. Mostre que P é invertível.

(a) Se P é invertível, é equivalente por linhas à matriz identidade I. Pelo Problema 6.19, $\{v_i\}$ e $\{u_i\}$ geram o mesmo subespaço de V. Assim, um é linearmente independente se, e só se, o outro é linearmente independente.

(b) Se P não é invertível, é equivalente por linhas a uma matriz com uma linha nula. Isso significa que $\{v_i\}$ gera um subespaço que tem um conjunto gerador com menos do que n elementos. Assim, $\{v_i\}$ é linearmente dependente.

(c) Essa afirmação é uma contraposição da afirmação (b), portanto segue de (b).

6.22 Demonstre o Teorema 6.6. Seja P a matriz de mudança de base de uma base S para uma base S' de um espaço vetorial V. Então, dado um vetor $v \in V$, qualquer, temos $P[v]_{S'} = [v]_S$, e, portanto, $P^{-1}[v]_S = [v]_{S'}$.

Sejam $S = \{u_1, ..., u_n\}$ e $S' = \{w_1, \ldots, w_n\}$, e suponha que, para $i = 1, ..., n$,

$$w_i = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j$$

Então P é a matriz quadrada de ordem n cuja j-ésima linha é

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Também suponha que $v = k_1w_1 + k_2w_2 + \cdots + k_nw_n = \sum_{i=1}^{n} k_iw_i$. Então

$$[v]_{S'} = [k_1, k_2, \ldots, k_n]^T \tag{2}$$

Substituindo os w_i nesta equação de v, obtemos

$$v = \sum_{i=1}^{n} k_iw_i = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}u_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)u_j$$
$$= \sum_{j=1}^{n}(a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)u_j$$

Dessa forma, $[v]_S$ é o vetor coluna cuja j-ésima entrada é

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \tag{3}$$

Por outro lado, a j-ésima entrada de $P[v]_{S'}$ é obtida multiplicando a j-ésima linha de P por $[v]_{S'}$, ou seja, (1) por (2). Contudo, o produto de (1) por (2) é (3). Assim, $P[v]_{S'}$ e $[v]_S$ têm as mesmas entradas. Assim, $P[v]_{S'} = [v]_{S'}$, como queríamos mostrar.

Além disso, a multiplicação da expressão obtida por P^{-1}, fornece $P^{-1}[v]_S = P^{-1}P[v]_{S'} = [v]_{S'}$.

Operadores lineares e mudança de base

6.23 Considere o operador linear F de $\mathbf{R}^2$ definido por $F(x, y) = (5x - y,\ 2x + y)$ e as bases de $\mathbf{R}^2$ a seguir.

$$E = \{e_1, e_2\} = \{(1, 0),\ (0, 1)\} \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1, 4),\ (2, 7)\}$$

(a) Encontre a matriz P de mudança de base de E para S e a matriz Q de mudança de base de S de volta para E.

(b) Encontre a matriz A que representa F na base E.

(c) Encontre a matriz B que representa F na base S.

(a) Como E é a base canônica, simplesmente escrevemos os vetores de S como colunas para obter a matriz P de mudança de base. Além disso, sabemos que $Q = P^{-1}$. Assim,

$$P = \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \qquad \text{e} \qquad Q = P^{-1} = \begin{bmatrix} -7 & 2 \\ 4 & -1 \end{bmatrix}$$

(b) Escrevemos os coeficientes de x e y em $F(x, y) = (5x - y,\ 2x + y)$ como linhas para obter

$$A = \begin{bmatrix} 5 & -1 \\ 2 & 1 \end{bmatrix}$$

(c) **Método 1** Encontramos as coordenadas de $F(u_1)$ e $F(u_2)$ em relação à base S. Isso pode ser feito calculando, primeiro, as coordenadas de um vetor (a, b) arbitrário de $\mathbf{R}^2$ em relação à base S. Temos

$$(a, b) = x(1, 4) + y(2, 7) = (x + 2y,\ 4x + 7y), \quad \text{e, portanto,} \quad \begin{aligned} x + 2y &= a \\ 4x + 7y &= b \end{aligned}$$

Resolvemos x e y em termos de a e b para obter $x = -7a + 2b,\ \ y = 4a - b$. Então

$$(a, b) = (-7a + 2b)u_1 + (4a - b)u_2$$

Agora usamos a fórmula de (a, b) para obter

$$\begin{aligned} F(u_1) &= F(1, 4) = (1, 6) \ = 5u_1 - 2u_2 \\ F(u_2) &= F(2, 7) = (3, 11) = \ u_1 + \ u_2 \end{aligned} \quad \text{e, portanto,} \quad B = \begin{bmatrix} 5 & 1 \\ -2 & 1 \end{bmatrix}$$

Método 2 Pelo Teorema 6.7, temos $B = P^{-1}AP$. Assim,

$$B = P^{-1}AP = \begin{bmatrix} -7 & 2 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ -2 & 1 \end{bmatrix}$$

6.24 Seja $A = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}$. Encontre a matriz B que representa o operador linear A em relação à base $S = \{u_1, u_2\} = \{[1, 3]^T,\ [2, 5]^T\}$. [Vimos que A define um operador linear $A\colon \mathbf{R}^2 \to \mathbf{R}^2$ em relação à base canônica E de $\mathbf{R}^2$.]

Método 1 Encontramos as coordenadas de $A(u_1)$ e $A(u_2)$ em relação à base S calculando, primeiro, as coordenadas de um vetor $[a, b]^T$ arbitrário de $\mathbf{R}^2$ em relação à base S. Pelo Problema 6.2,

$$[a, b]^T = (-5a + 2b)u_1 + (3a - b)u_2$$

Usando a fórmula de $[a, b]^T$, obtemos

$$A(u_1) = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 1 \end{bmatrix} = -53u_1 + 32u_2$$

e

$$A(u_2) = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 19 \\ 3 \end{bmatrix} = -89u_1 + 54u_2$$

Assim,

$$B = \begin{bmatrix} -53 & -89 \\ 32 & 54 \end{bmatrix}$$

Método 2 Usamos $B = P^{-1}AP$, onde P é a matriz de mudança de base da base canônica E para S. Para isso, simplesmente escrevemos os vetores de S como colunas para obter a matriz P de mudança de base e então usamos a fórmula de P^{-1}. Dessa forma,

$$P = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \quad \text{e} \quad P^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

Assim, $$B = P^{-1}AP = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -53 & -89 \\ 32 & 54 \end{bmatrix}$$

6.25 Seja $A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & -4 \\ 1 & -2 & 2 \end{bmatrix}$. Encontre a matriz B que representa o operador linear A em relação à base

$$S = \{u_1, u_2, u_3\} = \{[1, 1, 0]^T, \quad [0, 1, 1]^T, \quad [1, 2, 2]^T\}$$

[Vimos que A define um operador linear $A: \mathbf{R}^3 \to \mathbf{R}^3$ em relação à base canônica E de $\mathbf{R}^3$.]

Método 1 Encontramos as coordenadas de $A(u_1)$, $A(u_2)$, $A(u_3)$ em relação à base S calculando, primeiro, as coordenadas de um vetor $v = [a, b, c]^T$ arbitrário de $\mathbf{R}^3$ em relação à base S. Pelo Problema 6.16,

$$[v]_S = (b - c)u_1 + (-2a + 2b - c)u_2 + (a - b + c)u_3$$

Usando essa fórmula de $[a, b, c]^T$, obtemos

$$A(u_1) = [4, 7, -1]^T = 8u_1 + 7u_2 - 5u_3, \qquad A(u_2) = [4, 1, 0]^T = u_1 - 6u_2 + 3u_3$$
$$A(u_3) = [9, 4, 1]^T = 3u_1 - 11u_2 + 6u_3$$

Escrevendo os coeficientes de u_1, u_2, u_3 como colunas, obtemos

$$B = \begin{bmatrix} 8 & 1 & 3 \\ 7 & -6 & -11 \\ -5 & 3 & 6 \end{bmatrix}$$

Método 2 Usamos $B = P^{-1}AP$, onde P é a matriz de mudança de base da base canônica E para S. A matriz P (cujas colunas são, simplesmente, os vetores de S) e P^{-1} foram calculadas no Problema 6.16. Assim,

$$B = P^{-1}AP = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & -4 \\ 1 & -2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 1 & 3 \\ 7 & -6 & -11 \\ -5 & 3 & 6 \end{bmatrix}$$

6.26 Demonstre o Teorema 6.7. Seja P a matriz de mudança de base de uma base S para uma base S' de um espaço vetorial V. Então, dado um operador linear T qualquer de V, $[T]_{S'} = P^{-1}[T]_S P$.

Seja v um vetor de V. Então, pelo Teorema 6.6, $P[v]_{S'} = [v]_S$ e, portanto,

$$P^{-1}[T]_S P[v]_{S'} = P^{-1}[T]_S[v]_S = P^{-1}[T(v)]_S = [T(v)]_{S'}$$

Mas $[T]_{S'}[v]_{S'} = [T(v)]_{S'}$. Logo,

$$P^{-1}[T]_S P[v]_{S'} = [T]_{S'}[v]_{S'}$$

Como a aplicação $v \mapsto [v]_{S'}$ é sobre K^n, temos $P^{-1}[T]_S PX = [T]_{S'}X$, para cada $X \in K^n$. Assim, $P^{-1}[T]_S P = [T]_{S'}$, como queríamos mostrar.

Semelhança de matrizes

6.27 Sejam $A = \begin{bmatrix} 4 & -2 \\ 3 & 6 \end{bmatrix}$ e $P = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

(a) Encontre $B = P^{-1}AP$. (b) Verifique $\text{tr}(B) = \text{tr}(A)$. (c) Verifique $\det(B) = \det(A)$.

(a) Inicialmente calculamos P^{-1} usando a fórmula da inversa de uma matriz 2×2. Temos

$$P^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

Então

$$B = P^{-1}AP = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 25 & 30 \\ -\frac{27}{2} & -15 \end{bmatrix}$$

(b) $\text{tr}(A) = 4 + 6 = 10$ e $\text{tr}(B) = 25 - 15 = 10$. Logo, tr($B$) = tr($A$).

(c) $\det(A) = 24 + 6 = 30$ e $\det(B) = -375 + 405 = 30$. Logo, det($B$) = det($A$).

6.28 Encontre o traço de cada um dos operadores F de $\mathbf{R}^3$ do Problema 6.4.

Calculamos o traço (soma dos elementos diagonais) de qualquer representação matricial de F, por exemplo, a representação matricial $[F] = [F]_E$ de F em relação à base canônica E, dada no Problema 6.4

(a) $\text{tr}(F) = \text{tr}([F]) = 1 - 5 + 9 = 5$.

(b) $\text{tr}(F) = \text{tr}([F]) = 1 + 3 + 5 = 9$.

(c) $\text{tr}(F) = \text{tr}([F]) = 1 + 4 + 7 = 12$.

6.29 Escrevamos $A \approx B$ se A for similar a B, ou seja, se existir uma matriz invertível P tal que $A = P^{-1}BP$. Prove que $\approx$ é uma relação de equivalência (nas matrizes quadradas), ou seja,

(a) $A \approx A$, para cada A, (b) Se $A \approx B$, então $B \approx A$.

(c) Se $A \approx B$ e $B \approx C$, então $A \approx C$.

(a) A matriz identidade I é invertível e $I^{-1} = I$. Como $A = I^{-1}AI$, temos $A \approx A$.

(b) Como $A \approx B$, existe uma matriz P invertível tal que $A = P^{-1}BP$. Logo, $B = PAP^{-1} = (P^{-1})^{-1}AP$ e P^{-1} também é invertível. Assim, $B \approx A$.

(c) Como $A \approx B$, existe uma matriz P invertível tal que $A = P^{-1}BP$, e, como $B \approx C$, existe uma matriz Q invertível tal que $B = Q^{-1}CQ$. Logo,

$$A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P = (P^{-1}Q^{-1})C(QP) = (QP)^{-1}C(QP)$$

e QP também é invertível. Assim, $A \approx C$.

6.30 Seja B uma matriz semelhante a A, digamos, $B = P^{-1}AP$. Demonstre.

(a) $B^n = P^{-1}A^nP$, portanto, B^n é semelhante a A^n.

(b) $f(B) = P^{-1}f(A)P$, para qualquer polinômio $f(x)$, portanto, $f(B)$ é semelhante a $f(A)$.

(c) B é uma raiz de um polinômio $g(x)$ se, e só se, A é uma raiz de $g(x)$.

(a) A demonstração é por indução em n. Por hipótese, o resultado é válido para $n = 1$. Suponha que $n > 1$ e que o resultado seja válido para $n - 1$. Então

$$B^n = BB^{n-1} = (P^{-1}AP)(P^{-1}A^{n-1}P) = P^{-1}A^nP$$

(b) Suponha que $f(x) = a_nx^n + \cdots + a_1x + a_0$. Usando a distributividade à esquerda e à direita, bem como a parte (a), obtemos

$$\begin{aligned} P^{-1}f(A)P &= P^{-1}(a_nA^n + \cdots + a_1A + a_0I)P \\ &= P^{-1}(a_nA^n)P + \cdots + P^{-1}(a_1A)P + P^{-1}(a_0I)P \\ &= a_n(P^{-1}A^nP) + \cdots + a_1(P^{-1}AP) + a_0(P^{-1}IP) \\ &= a_nB^n + \cdots + a_1B + a_0I = f(B) \end{aligned}$$

(c) Pela parte (b), $g(B) = 0$ se, e só se, $P^{-1}g(A)P = 0$ se, e só se, $g(A) = P0P^{-1} = 0$.

Representação matricial de transformações lineares gerais

6.31 Seja $F\colon \mathbf{R}^3 \to \mathbf{R}^2$ a transformação linear definida por $F(x, y, z) = (3x + 2y - 4z,\ x - 5y + 3z)$.

(a) Encontre a matriz de F nas bases de $\mathbf{R}^3$ e $\mathbf{R}^2$ dadas a seguir.

$S = \{w_1, w_2, w_3\} = \{(1, 1, 1),\ (1, 1, 0),\ (1, 0, 0)\}$ e $S' = \{u_1, u_2\} = \{(1, 3),\ (2, 5)\}$

(b) Verifique o Teorema 6.10, a saber, que a ação de F é preservada pela representação matricial de F. Em outras palavras, que, dado v de $\mathbf{R}^3$ qualquer, temos $[F]_{S,S'}[v]_S = [F(v)]_{S'}$.

(a) Pelo Problema 6.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Assim,

$$\begin{aligned} F(w_1) &= F(1,1,1) = (1,-1) = -7u_1 + 4u_2 \\ F(w_2) &= F(1,1,0) = (5,-4) = -33u_1 + 19u_2 \\ F(w_3) &= F(1,0,0) = (3,1) = -13u_1 + 8u_2 \end{aligned}$$

Escrevemos as coordenadas de $F(w_1)$, $F(w_2)$, $F(w_3)$ como colunas para obter

$$[F]_{S,S'} = \begin{bmatrix} -7 & -33 & 13 \\ 4 & 19 & 8 \end{bmatrix}$$

(b) Se $v = (x, y\ z)$, então, pelo Problema 6.5, $v = zw_1 + (y - z)w_2 + (x - y)w_3$. Também

$$F(v) = (3x + 2y - 4z,\ \ x - 5y + 3z) = (-13x - 20y + 26z)u_1 + (8x + 11y - 15z)u_2$$

Logo, $\quad [v]_S = (z,\ y - z,\ x - y)^T \quad$ e $\quad [F(v)]_{S'} = \begin{bmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{bmatrix}$

Assim, $\quad [F]_{S,S'}[v]_S = \begin{bmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{bmatrix} \begin{bmatrix} z \\ y - x \\ x - y \end{bmatrix} = \begin{bmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{bmatrix} = [F(v)]_{S'}$

6.32 Seja $F\colon \mathbf{R}^n \to \mathbf{R}^m$ a transformação linear definida como segue.

$$F(x_1, x_2, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ \ a_{21}x_1 + \cdots + a_{2n}x_n,\ \ \ldots,\ \ a_{m1}x_1 + \cdots + a_{mn}x_n)$$

(a) Mostre que as linhas da matriz $[F]$ que representa F em relação às bases canônicas de $\mathbf{R}^n$ e $\mathbf{R}^m$ são os coeficiente dos x_i nos componentes de $F(x_1, ..., x_n)$.

(b) Encontre a representação matricial de cada uma das transformações lineares seguintes em relação às bases canônicas dos $\mathbf{R}^n$.

(i) $F\colon \mathbf{R}^2 \to \mathbf{R}^3$ definida por $F(x, y) = (3x - y,\quad 2x + 4y,\quad 5x - 6y)$.
(ii) $F\colon \mathbf{R}^4 \to \mathbf{R}^2$ definida por $F(x, y, s, t) = (3x - 4y + 2s - 5t,\quad 5x + 7y - s - 2t)$.
(iii) $F\colon \mathbf{R}^3 \to \mathbf{R}^4$ definida por $F(x, y, z) = (2x + 3y - 8z,\quad x + y + z,\quad 4x - 5z,\quad 6y)$.

(a) Temos

$$\begin{aligned} F(1,0,\ldots,0) &= (a_{11}, a_{21}, \ldots, a_{m1}) \\ F(0,1,\ldots,0) &= (a_{12}, a_{22}, \ldots, a_{m2}) \\ &\cdots\cdots\cdots\cdots \\ F(0,0,\ldots,1) &= (a_{1n}, a_{2n}, \ldots, a_{mn}) \end{aligned} \quad \text{e, portanto,} \quad [F] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

(b) Pela parte (a), basta olhar para os coeficientes das incógnitas x, y em $F(x, y, ...)$. Assim,

$$\text{(i)}\quad [F] = \begin{bmatrix} 3 & -1 \\ 2 & 4 \\ 5 & -6 \end{bmatrix}, \qquad \text{(ii)}\quad [F] = \begin{bmatrix} 3 & -4 & 2 & -5 \\ 5 & 7 & -1 & -2 \end{bmatrix}, \qquad \text{(iii)}\quad [F] = \begin{bmatrix} 2 & 3 & -8 \\ 1 & 1 & 1 \\ 4 & 0 & -5 \\ 0 & 6 & 0 \end{bmatrix}$$

6.33 Seja $A = \begin{bmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{bmatrix}$. Vimos que A determina uma transformação linear $F\colon \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(v) = Av$, em que os vetores são escritos como colunas. Encontre a matriz $[F]$ que representa a transformação em relação às bases de $\mathbf{R}^3$ e $\mathbf{R}^2$ dadas a seguir.

(a) As bases canônicas de $\mathbf{R}^3$ e $\mathbf{R}^2$.

(b) $S = \{w_1, w_2, w_3\} = \{(1,1,1),\ (1,1,0),\ (1,0,0)\}$ e $S' = \{u_1, u_2\} = \{(1,3),\ (2,5)\}$.

(a) Em relação às bases canônicas, $[F]$ é a matriz A.

(b) Pelo Problema 6.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Assim,

$$F(w_1) = \begin{bmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = -12u_1 + 8u_2$$

$$F(w_2) = \begin{bmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ -3 \end{bmatrix} = -41u_1 + 24u_2$$

$$F(w_3) = \begin{bmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -8u_1 + 5u_2$$

Escrevendo as coordenadas de $F(w_1)$, $F(w_2)$, $F(w_3)$ como colunas, obtemos $[F] = \begin{bmatrix} -12 & -41 & -8 \\ 8 & 24 & 5 \end{bmatrix}$.

6.34 Considere a transformação linear T de $\mathbf{R}^2$ definida por $T(x, y) = (2x - 3y, \quad x + 4y)$ e as bases de $\mathbf{R}^2$ a seguir.

$$E = \{e_1, e_2\} = \{(1, 0), \ (0, 1)\} \qquad \text{e} \qquad S = \{u_1, u_2\} = \{(1, 3), \ (2, 5)\}$$

(a) Encontre a matriz A que representa T nas bases E e S.

(b) Encontre a matriz B que representa T nas bases S e E.

(Podemos interpretar T como uma transformação linear de um espaço em outro, cada um com sua base.)

(a) Pelo Problema 6.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Logo,

$$\begin{aligned} T(e_1) &= T(1, 0) = (2, 1) &&= -8u_1 + 5u_2 \\ T(e_2) &= T(0, 1) = (-3, 4) &&= 23u_1 - 13u_2 \end{aligned} \qquad \text{e, portanto,} \quad A = \begin{bmatrix} -8 & 23 \\ 5 & -13 \end{bmatrix}$$

(b) Temos

$$\begin{aligned} T(u_1) &= T(1, 3) = (-7, 13) &&= -7e_1 + 13e_2 \\ T(u_2) &= T(2, 5) = (-11, 22) &&= -11e_1 + 22e_2 \end{aligned} \qquad \text{e, portanto,} \quad B = \begin{bmatrix} -7 & -11 \\ 13 & 22 \end{bmatrix}$$

6.35 Qual é a relação entre as matrizes A e B do Problema 6.34?

Pelo Teorema 6.12, as matrizes A e B são equivalentes, ou seja, existem matrizes P e Q não singulares tais que $B = Q^{-1}AP$, sendo P a matriz de mudança de base de S para E e Q a matriz de mudança de base de E para S. Assim,

$$P = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}, \qquad Q = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}, \qquad Q^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$$

e

$$Q^{-1}AP = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -8 & -23 \\ 5 & -13 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} -7 & -11 \\ 13 & 22 \end{bmatrix} = B$$

6.36 Demonstre o Teorema 6.14. Seja $F\colon V \to U$ uma transformação linear, digamos, com pos(F) = r. Então existem bases de V e de U tais que a representação matricial de F é da forma seguinte, em que I_r é a matriz identidade de ordem r.

$$A = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Sejam dim $V = m$ e dim $U = n$. Sejam W o núcleo de F e U' a imagem de F. É dado que pos(F) = r. Logo, a dimensão do núcleo de F é $m - r$. Tomando uma base $\{w_1, \ldots, w_{m-r}\}$ do núcleo de F, estendemos essa base a uma base de V, como segue.

$$\{v_1, \ldots, v_r, w_1, \ldots, w_{m-r}\}$$

Sejam

$$u_1 = F(v_1), \ u_2 = F(v_2), \ \ldots, \ u_r = F(v_r)$$

Então $\{u_1, ..., u_r\}$ é uma base de U', a imagem de F. Estendemos essa base a uma base de U, digamos,

$$\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$$

Observe que

$$\begin{aligned}
F(v_1) &= u_1 = 1u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
F(v_2) &= u_2 = 0u_1 + 1u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
F(v_r) &= u_r = 0u_1 + 0u_2 + \cdots + 1u_r + 0u_{r+1} + \cdots + 0u_n \\
F(w_1) &= 0 \;\;= 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
F(w_{m-r}) &= 0 \;\;= 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n
\end{aligned}$$

Assim, a matriz de F nessas bases tem o formato exigido.

Problemas Complementares

Matrizes e operadores lineares

6.37 Seja $F: \mathbf{R}^2 \to \mathbf{R}^2$ definido por $F(x, y) = (4x + 5y,\ 2x - y)$.

(a) Encontre a matriz A que representa F na base canônica E.

(b) Encontre a matriz B que representa F na base $S = \{u_1, u_2\} = \{(1, 4),\ (2, 9)\}$.

(c) Encontre P tal que $B = P^{-1}AP$.

(d) Para $v = (a, b)$, encontre $[v]_S$ e $[F(v)]_S$. Verifique que $[F]_S[v]_S = [F(v)]_S$.

6.38 Seja $A: \mathbf{R}^2 \to \mathbf{R}^2$ definido pela matriz $A = \begin{bmatrix} 5 & -1 \\ 2 & 4 \end{bmatrix}$.

(a) Encontre a matriz B que representa A em relação à base $S = \{u_1, u_2\} = \{(1, 3),\ (2, 8)\}$. (Lembre que A representa um operador em relação à base canônica E.)

(b) Para $v = (a, b)$, encontre $[v]_S$ e $[A(v)]_S$.

6.39 Para cada operador L de $\mathbf{R}^2$ dado, encontre a matriz A que representa L (em relação à base canônica de $\mathbf{R}^2$).

(a) L é a rotação de $\mathbf{R}^2$ no sentido anti-horário por $45°$.

(b) L é a reflexão de $\mathbf{R}^2$ em torno da reta $y = x$.

(c) L é definido por $L(1, 0) = (3, 5)$ e $L(0, 1) = (7, -2)$.

(d) L é definido por $L(1, 1) = (3, 7)$ e $L(1, 2) = (5, -4)$.

6.40 Encontre a representação matricial de cada operador T de $\mathbf{R}^3$ dado em relação à base canônica de $\mathbf{R}^3$

(a) $T(x, y, z) = (x, y, 0)$. (b) $T(x, y, z) = (z,\ y + z,\ x + y + z)$.

(c) $T(x, y, z) = (2x - 7y - 4z,\ 3x + y + 4z,\ 6x - 8y + z)$.

6.41 Repita o Problema 6.40 usando a base $S = \{u_1, u_2, u_3\} = \{(1, 1, 0),\ (1, 2, 3),\ (1, 3, 5)\}$.

6.42 Seja L o operador de $\mathbf{R}^3$ definido por

$$L(1, 0, 0) = (1, 1, 1), \qquad L(0, 1, 0) = (1, 3, 5), \qquad L(0, 0, 1) = (2, 2, 2)$$

(a) Encontre a matriz A que representa L em relação à base canônica E.

(b) Encontre a matriz B que representa L em relação à base S do Problema 6.41.

6.43 Seja $\mathbf{D}$ o operador derivada, ou seja, $\mathbf{D}(f(t)) = df/dt$. Cada um dos conjuntos a seguir é uma base de um espaço vetorial V de funções. Encontre a representação matricial de $\mathbf{D}$ em cada base.

(a) $\{e^t, e^{2t}, te^{2t}\}$. (b) $\{1, t, \operatorname{sen} 3t, \cos 3t\}$. (c) $\{e^{5t}, te^{5t}, t^2e^{5t}\}$.

6.44 Seja **D** o operador derivada do espaço vetorial V de funções de base $S = \{\text{sen}\,\theta, \cos\theta\}$.

(a) Encontre a matriz $A = [\mathbf{D}]_S$. (b) Use A para mostrar que **D** é um zero de $f(t) = t^2 + 1$.

6.45 Seja V o espaço vetorial das matrizes 2×2. Considere a matriz M a seguir e a base canônica E de V.

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \text{e} \qquad E = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Encontre a representação matricial de cada um dos operadores T de V a seguir em relação a E.

(a) $T(A) = MA$, (b) $T(A) = AM$, (c) $T(A) = MA - AM$.

6.46 Sejam $\mathbf{1}_V$ e $\mathbf{0}_V$ os operadores identidade e nulo, respectivamente, de um espaço vetorial V. Mostre que, dada qualquer base S de V, (a) $[\mathbf{1}_V]_S = I$, a matriz identidade, (b) $[\mathbf{0}_V]_S = 0$, a matriz nula.

Mudança de base

6.47 Encontre a matriz P de mudança de base da base canônica E de $\mathbf{R}^2$ para a base S, a matriz Q de mudança de base de S de volta para E e as coordenadas de $v = (a, b)$ em relação a S, nos casos dados a seguir.

(a) $S = \{(1, 2),\ (3, 5)\}$. (c) $S = \{(2, 5),\ (3, 7)\}$.
(b) $S = \{(1, -3),\ (3, -8)\}$. (d) $S = \{(2, 3),\ (4, 5)\}$.

6.48 Considere as bases $S = \{(1, 2),\ (2, 3)\}$ e $S' = \{(1, 3),\ (1, 4)\}$ de $\mathbf{R}^2$. Encontre a matriz de mudança de base

(a) P de S para S'. (b) Q de S' de volta para S.

6.49 Suponha que os eixos x e y do plano $\mathbf{R}^2$ tenham sido girados por 30° no sentido anti-horário para formar novos eixos x' e y' do plano. Encontre

(a) os vetores unitários na direção dos novos eixos x' e y',

(b) a matriz P de mudança de base para o novo sistema,

(b) as novas coordenadas dos pontos $A(1, 3)$, $B(2, -5)$, $C(a, b)$.

6.50 Encontre a matriz P de mudança de base da base canônica E de $\mathbf{R}^3$ para a base S, a matriz Q de mudança de base de S de volta para E e as coordenadas de $v = (a, b, c)$ em relação a S, nos casos de S dados a seguir.

(a) $u_1 = (1, 1, 0), u_2 = (0, 1, 2), u_3 = (0, 1, 1)$.
(b) $u_1 = (1, 0, 1), u_2 = (1, 1, 2), u_3 = (1, 2, 4)$.
(c) $u_1 = (1, 2, 1), u_2 = (1, 3, 4), u_3 = (2, 5, 6)$.

6.51 Suponha que S_1, S_2, S_3 sejam bases de V. Sejam P e Q as matrizes de mudança de base, respectivamente, de S_1 para S_2 e de S_2 para S_3. Prove que PQ é a matriz de mudança de base de S_1 para S_3.

Operadores lineares e mudança de base

6.52 Considere o operador linear de $\mathbf{R}^2$ definido por $F(x, y) = (5x + y,\ 3x - 2y)$ e as bases de $\mathbf{R}^2$ a seguir.

$$S = \{(1, 2),\ (2, 3)\} \qquad \text{e} \qquad S' = \{(1, 3),\ (1, 4)\}$$

(a) Encontre a matriz A que representa F em relação à base S.

(b) Encontre a matriz B que representa F em relação à base S'.

(c) Encontre a matriz P de mudança de base de S para S'.

(d) Qual é a relação entre A e B?

6.53 Seja $A\colon \mathbf{R}^2 \to \mathbf{R}^2$ definido pela matriz $A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}$. Encontre a matriz B que representa o operador linear A em relação a cada uma das bases dadas. (a) $S = \{(1, 3)^T,\ (2, 5)^T\}$. (b) $S = \{(1, 3)^T,\ (2, 4)^T\}$.

6.54 Seja $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ definido por $F(x,y) = (x-3y,\quad 2x-4y)$. Encontre a matriz A que representa F em relação a cada uma das bases dadas. (a) $S = \{(2,5),\ (3,7)\}$. (b) $S = \{(2,3),\ (4,5)\}$.

6.55 Seja $A\colon \mathbf{R}^3 \to \mathbf{R}^3$ definido pela matriz $A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 7 & 4 \\ 1 & 4 & 3 \end{bmatrix}$. Encontre a matriz B que representa o operador linear A em relação à base $S = \{(1,1,1)^T,\ (0,1,1)^T,\ (1,2,3)^T\}$.

Semelhança de matrizes

6.56 Sejam $A = \begin{bmatrix} 1 & 1 \\ 2 & -3 \end{bmatrix}$ e $P = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix}$.

(a) Encontre $B = P^{-1}AP$. (b) Verifique $\operatorname{tr}(B) = \operatorname{tr}(A)$. (c) Verifique $\det(B) = \det(A)$.

6.57 Encontre o traço e o determinante de cada um dos operadores de $\mathbf{R}^2$ dados.

(a) $F(x,y) = (2x-3y,\ 5x+4y)$. (b) $G(x,y) = (ax+by,\ cx+dy)$.

6.58 Encontre o traço e o determinante de cada um dos operadores de $\mathbf{R}^3$ dados.

(a) $F(x,y,z) = (x+3y,\ 3x-2z,\ x-4y-3z)$.

(b) $G(x,y,z) = (y+3z,\ 2x-4z,\ 5x+7y)$.

6.59 Sejam $S = \{u_1, u_2\}$ uma base de V e $T\colon V \to V$ o operador definido por $T(u_1) = 3u_1 - 2u_2$ e $T(u_2) = u_1 + 4u_2$. Seja $S' = \{w_1, w_2\}$ a base de V dada por $w_1 = u_1 + u_2$ e $w_2 = 2u_1 + 3u_2$.

(a) Encontre as matrizes A e B que representam T em relação às bases S e S', respectivamente.

(b) Encontre a matriz P tal que $B = P^{-1}AP$.

6.60 Seja A uma matriz 2×2 tal que somente A é semelhante a A. Mostre que A é uma matriz escalar, ou seja, $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$.

6.61 Mostre que todas as matrizes semelhantes a uma matriz invertível são invertíveis. Mais geralmente, mostre que matrizes semelhantes têm o mesmo posto.

Representação matricial de transformações lineares gerais

6.62 Encontre a representação matricial de cada uma das transformações lineares em relação às bases canônicas dos $\mathbf{R}^n$.

(a) $F\colon \mathbf{R}^3 \to \mathbf{R}^2$, definida por $F(x,y,z) = (2x-4y+9z,\ 5x+3y-2z)$.

(b) $F\colon \mathbf{R}^2 \to \mathbf{R}^4$, definida por $F(x,y) = (3x+4y,\ 5x-2y,\ x+7y,\ 4x)$.

(c) $F\colon \mathbf{R}^4 \to \mathbf{R}$, definida por $F(x_1,x_2,x_3,x_4) = 2x_1 + x_2 - 7x_3 - x_4$.

6.63 Seja $G\colon \mathbf{R}^3 \to \mathbf{R}^2$ definida por $G(x,y,z) = (2x+3y-z,\ 4x-y+2z)$.

(a) Encontre a matriz A que representa G em relação às bases

$$S = \{(1,1,0),\ (1,2,3),\ (1,3,5)\} \qquad \text{e} \qquad S' = \{(1,2),\ (2,3)\}$$

(b) Para $v = (a,b,c)$ de $\mathbf{R}^3$, encontre $[v]_S$ e $[G(v)]_{S'}$. (c) Verifique que $A[v]_S = [G(v)]_{S'}$.

6.64 Seja $H\colon \mathbf{R}^2 \to \mathbf{R}^2$ definida por $H(x,y) = (2x+7y,\ x-3y)$ e considere as bases de $\mathbf{R}^2$ a seguir.

$$S = \{(1,1),\ (1,2)\} \qquad \text{e} \qquad S' = \{(1,4),\ (1,5)\}$$

(a) Encontre a matriz A que representa H em relação às bases S e S'.

(b) Encontre a matriz B que representa H em relação às bases S' e S.

6.65 Seja $F\colon \mathbf{R}^3 \to \mathbf{R}^2$ definido por $F(x,y,z) = (2x+y-z, \quad 3x-2y+4z)$.

(a) Encontre a matriz A que representa F em relação às bases

$$S = \{(1,1,1), \quad (1,1,0), \quad (1,0,0)\} \qquad \text{e} \qquad S' = (1,3), \quad (1,4)\}$$

(b) Verifique que $A[v]_{S'} = [F(v)]_S$, para qualquer $v = (a, b, c)$ de $\mathbf{R}^3$

6.66 Sejam S e S' bases de V e seja $\mathbf{1}_V$ a transformação identidade de V. Mostre que a matriz A que representa $\mathbf{1}_v$ em relação às bases S e S' é a inversa da matriz P de mudança de base de S para S', ou seja, $A = P^{-1}$.

6.67 Demonstre os Teoremas (a) 6.10, (b) 6.11, (c) 6.12, (d) 6.13. [*Sugestão*: ver as demonstrações análogas dos Teoremas 6.1, 6.2, 6.3 e 6.7 nos Problemas 6.9, 6.10, 6.11 e 6.26, respectivamente.]

Problemas variados

6.68 Suponha que $F\colon V \to V$ seja linear. Dizemos que um subespaço W de V é *invariante* por F se $F(W) \subseteq W$. Suponha que W seja invariante por F e que dim $W = r$. Mostre que F tem uma representação em matriz triangular superior em blocos $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, em que A é uma submatriz quadrada de ordem r.

6.69 Suponha que $V = U + W$ e que U e W sejam, ambos, invariantes por um operador linear $F\colon V \to V$. Suponha, também, que dim $U = r$ e que dim $W = s$. Mostre que F tem uma representação matricial $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, em que A e B são submatrizes quadradas de ordens r e s, respectivamente.

6.70 Dizemos que dois operadores lineares F e G de V são *semelhantes* se existir um operador linear invertível T de V tal que $G = T^{-1} \circ F \circ T$. Prove as afirmações seguintes.

(a) F e G são semelhantes se, e só se, as matrizes $[F]_S$ e $[G]_S$ são semelhantes, para qualquer base S de V.

(b) Se F é diagonalizável (ou seja, semelhante a uma matriz diagonal), então qualquer matriz G semelhante a F também é diagonalizável.

Respostas dos Problemas Complementares

Notação: $M = [R_1; \quad R_2; \quad \ldots]$ denota uma matriz M de linhas R_1, R_2, ...

6.37 (a) $A = [4,5; \quad 2,-1]$; (b) $B = [220,487; \quad -98,-217]$; (c) $P = [1,2; \quad 4,9]$;
(d) $[v]_S = [9a-2b, \quad -4a+b]^T$ e $[F(v)]_S = [32a+47b, \quad -14a-21b]^T$

6.38 (a) $B = [-6,-28; \quad 4,15]$;
(b) $[v]_S = [4a-b, -\frac{3}{2}a+\frac{1}{2}b]^T$ e $[A(v)]_S = [18a-8b, \quad \frac{1}{2}(-13a+7b)]$

6.39 (a) $[\sqrt{2},-\sqrt{2}; \quad \sqrt{2},\sqrt{2}]$; (b) $[0,1; \quad 1,0]$; (c) $[3,7; \quad 5,-2]$;
(d) $[1,2; \quad 18,-11]$

6.40 (a) $[1,0,0; \quad 0,1,0; \quad 0,0,0]$; (b) $[0,0,1; \quad 0,1,1; \quad 1,1,1]$;
(c) $[2,-7,-4; \quad 3,1,4; \quad 6,-8,1]$

6.41 (a) $[1,3,5; \quad 0,-5,-10; \quad 0,3,6]$; (b) $[0,1,2; \quad -1,2,3; \quad 1,0,0]$;
(c) $[15,65,104; \quad -49,-219,-351; \quad 29,130,208]$

6.42 (a) $[1,1,2; \quad 1,3,2; \quad 1,5,2]$; (b) $[0; \quad 2,14,22; \quad 0,-5,-8]$

6.43 (a) $[1,0,0; \quad 0,2,1; \quad 0,0,2]$; (b) $[0,1,0,0; \quad 0; \quad 0,0,0,-3; \quad 0,0,3,0]$;
(c) $[5,1,0; \quad 0,5,2; \quad 0,0,5]$

6.44 (a) $A = [0, -1; \quad 1, 0]$; (b) $A^2 + I = 0$

6.45 (a) $[a, 0, b, 0; \quad 0, a, 0, b; \quad c, 0, d, 0; \quad 0, c, 0, d]$;
(b) $[a, c, 0, 0; \quad b, d, 0, 0; \quad 0, 0, a, c; \quad 0, 0, b, d]$;
(c) $[0, -c, b, 0; \quad -b, a-d, 0, b; \quad c, 0, d-a, -c; \quad 0, c, -b, 0]$

6.47 (a) $[1, 3; \quad 2, 5]$, $[-5, 3; \quad 2, -1]$, $[v] = [-5a + 3b, \quad 2a - b]^T$;
(b) $[1, 3; \quad -3, -8]$, $[-8, -3; \quad 3, 1]$, $[v] = [-8a - 3b, \quad 3a + b]^T$;
(c) $[2, 3; \quad 5, 7]$, $[-7, 3; \quad 5, -2]$, $[v] = [-7a + 3b, \quad 5a - 2b]^T$;
(d) $[2, 4; \quad 3, 5]$, $[-\frac{5}{2}, 2; \quad \frac{3}{2}, -1]$, $[v] = [-\frac{5}{2}a + 2b, \quad \frac{3}{2}a - b]^T$

6.48 (a) $P = [3, 5; \quad -1, -2]$; (b) $Q = [2, 5; \quad -1, -3]$

6.49 Aqui $K = \sqrt{3}$.

(a) $\frac{1}{2}(K, 1)$, $\frac{1}{2}(-1, K)$;
(b) $P = \frac{1}{2}[K, -1; \quad 1, K]$;
(c) $\frac{1}{2}[K + 3, 3K - 1]^T$, $\frac{1}{2}[2K - 5, -5K - 2]^T$, $\frac{1}{2}[aK + b, bK - a]^T$

6.50 P é a matriz cujas colunas são u_1, u_2, u_3, $Q = P^{-1}$, $[v] = Q[a, b, c]^T$.

(a) $Q = [1, 0, 0; \quad 1, -1, 1; \quad -2, 2, -1]$, $[v] = [a, \quad a - b + c, \quad -2a + 2b - c]^T$;
(b) $Q = [0, -2, 1; \quad 2, 3, -2; \quad -1, -1, 1]$, $[v] = [-2b + c, \quad 2a + 3b - 2c, \quad -a - b + c]^T$;
(c) $Q = [-2, 2, -1; \quad -7, 4, -1; \quad 5, -3, 1]$, $[v] = [-2a + 2b - c, \quad -7a + 4b - c, \quad 5a - 3b + c]^T$

6.52 (a) $[-23, -39; \quad 15, 26]$; (b) $[35, 41; \quad -27, -32]$; (c) $[3, 5; \quad -1, -2]$; (d) $B = P^{-1}AP$

6.53 (a) $[28, 47; \quad -15, -25]$; (b) $[13, 18; \quad -\frac{15}{2}, -10]$

6.54 (a) $[43, 60; \quad -33, -46]$; (b) $\frac{1}{2}[3, 7; \quad -5, -9]$

6.55 $[10, 8, 20; \quad 13, 11, 28; \quad -5, -4, -10]$

6.56 (a) $[-34, 57; \quad -19, 32]$; (b) $\text{tr}(B) = \text{tr}(A) = -2$; (c) $\det(B) = \det(A) = -5$

6.57 (a) $\text{tr}(F) = 6, \det(F) = 23$; (b) $\text{tr}(G) = a + d, \det(G) = ad - bc$

6.58 (a) $\text{tr}(F) = -2, \ \det(F) = 13$; (b) $\text{tr}(G) = 0, \ \det(G) = 22$

6.59 (a) $A = [3, 1; \quad -2, 4]$, $B = [8, 11; \quad -2, -1]$; (b) $P = [1, 2; \quad 1, 3]$

6.62 (a) $[2, -4, 9; \quad 5, 3, -2]$; (b) $[3, 5, 1, 4; \quad 4, -2, 7, 0]$; (c) $[2, 1, -7, -1]$

6.63 (a) $[-9, 1, 4; \quad 7, 2, 1]$; (b) $[v]_S = [-a + 2b - c, \quad 5a - 5b + 2c, \quad -3a + 3b - c]^T$, e $[G(v)]_{S'} = [2a - 11b + 7c, \quad 7b - 4c]^T$

6.64 (a) $A = [47, 85; \quad -38, -69]$; (b) $B = [71, 88; \quad -41, -51]$

6.65 $A = [3, 11, 5; \quad -1, -8, -3]$

Capítulo 7

Espaços com Produto Interno, Ortogonalidade

7.1 INTRODUÇÃO

A definição de um espaço vetorial V envolve um corpo arbitrário K. Neste capítulo restringimos K ao corpo real $\mathbf{R}$, caso em que V é dito *espaço vetorial real*; nas últimas seções deste capítulo estendemos nossos resultados para o caso em que K é o corpo complexo $\mathbf{C}$, caso em que dizemos que V é um *espaço vetorial complexo*. Também continuamos com nossa notação, na qual

u, v, w são vetores de V
a, b, c, k, r são escalares de K

Além disso, os espaços vetoriais V deste capítulo têm dimensão finita, salvo menção explícita em contrário.

Observe que os conceitos de "comprimento" e "ortogonalidade" não apareceram na investigação de espaços vetoriais arbitrários V (embora tenham aparecido na Seção 1.4 sobre os espaços $\mathbf{R}^n$ e $\mathbf{C}^n$). Agora acrescentamos uma estrutura adicional a um espaço vetorial V para obter um espaço com produto interno e, nesse contexto, aqueles conceitos estão definidos.

7.2 ESPAÇOS COM PRODUTO INTERNO

Começamos com uma definição.

DEFINIÇÃO Seja V um espaço vetorial real. Suponha que a cada par de vetores $u, v \in V$ esteja associado um número real, denotado por $\langle u, v\rangle$. Essa função é denominada *produto interno (real)* de V se satisfizer os axiomas seguintes.

[I_1] **(Linearidade)** $\langle au_1 + bu_2,\ v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$.
[I_2] **(Simetria)** $\langle u, v\rangle = \langle v, u\rangle$.
[I_3] **(Positividade)** $\langle u, u\rangle \geq 0.$; e $\langle u, u\rangle = 0$ se, e só se, $u = 0$.

O espaço vetorial V com um produto interno é denominado *espaço com produto interno (real)*.

O axioma [I_1] afirma que uma função produto interno é linear na primeira posição. Usando [I_1] e o axioma da simetria [I_2], obtemos

$$\langle u,\ cv_1 + dv_2\rangle = \langle cv_1 + dv_2,\ u\rangle = c\langle v_1, u\rangle + d\langle v_2, u\rangle = c\langle u, v_1\rangle + d\langle u, v_2\rangle$$

Isto é, a função produto interno também é linear na segunda posição. Combinando essas duas propriedades e usando indução, obtemos a fórmula geral seguinte.

$$\left\langle \sum_i a_i u_i,\ \sum_j b_j v_j \right\rangle = \sum_i \sum_j a_i b_j \langle u_i, v_j\rangle$$

Isso significa que o produto interno de uma combinação linear de vetores é igual à combinação linear dos produtos internos dos vetores.

Exemplo 7.1 Seja V um espaço com produto interno real. Então, por linearidade,

$$\begin{aligned}\langle 3u_1 - 4u_2,\ \ 2v_1 - 5v_2 + 6v_3\rangle &= 6\langle u_1, v_1\rangle - 15\langle u_1, v_2\rangle + 18\langle u_1, v_3\rangle \\ &\quad - 8\langle u_2, v_1\rangle + 20\langle u_2, v_2\rangle - 24\langle u_2, v_3\rangle\end{aligned}$$

$$\begin{aligned}\langle 2u - 5v,\ \ 4u + 6v\rangle &= 8\langle u, u\rangle + 12\langle u, v\rangle - 20\langle v, u\rangle - 30\langle v, v\rangle \\ &= 8\langle u, u\rangle - 8\langle v, u\rangle - 30\langle v, v\rangle\end{aligned}$$

Observe que, na última equação, usamos a simetria, que $\langle u, v\rangle = \langle v, u\rangle$.

OBSERVAÇÃO O axioma $[I_1]$ sozinho implica $\langle 0, 0\rangle = \langle 0v, 0\rangle = 0\langle v, 0\rangle = 0$. Assim, $[I_1]$, $[I_2]$ e $[I_3]$ são equivalentes a $[I_1]$ e $[I_2]$ juntos com o axioma seguinte.

$[I_3']$ Se $u \neq 0$, então $\langle u, u\rangle$ é positivo.

Ou seja, uma função que satisfaz $[I_1]$, $[I_2]$ e $[I_3']$ é um produto interno.

Norma de um vetor

Pelo terceiro axioma $[I_3]$ de um produto interno, $\langle u, u\rangle$ é não negativo, para qualquer vetor u e, portanto, existe a raiz quadrada não negativa desse valor. Usamos a notação

$$\|u\| = \sqrt{\langle u, u\rangle}$$

Esse número não negativo é denominado *norma* ou *comprimento* de u. A relação $\|u\|^2 = \langle u, u\rangle$ é muito utilizada.

OBSERVAÇÃO Se $\|u\| = 1$ ou, equivalentemente, se $\langle u, u\rangle = 1$, dizemos que u é um *vetor unitário* e que o vetor está *normalizado*. Um vetor não nulo v qualquer de V pode ser multiplicado pelo recíproco de seu comprimento para obter o vetor unitário

$$\hat{v} = \frac{1}{\|v\|}\, v$$

que é um múltiplo positivo de v. Esse processo é denominado *normalização* de v.

7.3 EXEMPLOS DE ESPAÇOS COM PRODUTO INTERNO

Nesta seção listamos os principais exemplos de espaços com produto interno utilizados neste texto.

O espaço euclidiano $\mathbf{R}^n$

Considere o espaço vetorial $\mathbf{R}^n$. O *produto escalar* de $\mathbf{R}^n$ é definido por

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

onde $u = (a_i)$ e $v = (b_i)$. Essa função define um produto interno em $\mathbf{R}^n$. A norma $\|u\|$ do vetor $u = (a_i)$ nesse espaço é dada a seguir.

$$\|u\| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

Por outro lado, pelo Teorema de Pitágoras, a distância da origem O de $\mathbf{R}^3$ ao ponto $P(a, b, c)$ é dada por $\sqrt{a^2 + b^2 + c^2}$. Essa é precisamente igual à norma, acima definida, do vetor $v = (a, b, c)$ de $\mathbf{R}^3$. Como o Teorema de Pitágoras é uma consequência dos axiomas da Geometria Euclidiana, o espaço vetorial $\mathbf{R}^n$ com o produto interno e a norma dados é denominado *espaço euclidiano n-dimensional*. Embora existam muitas maneiras de definir

produtos internos em $\mathbf{R}^n$, vamos sempre supor definido esse produto interno, denominado *produto interno canônico*, salvo menção explícita em contrário.

OBSERVAÇÃO Muitas vezes representamos os vetores de $\mathbf{R}^n$ como vetores coluna, ou seja, matrizes coluna $n \times 1$. Nesse caso, a fórmula

$$\langle u, v \rangle = u^T v$$

define o produto interno canônico de $\mathbf{R}^n$.

Exemplo 7.2 Sejam $u = (1, 3, -4, 2)$, $v = (4, -2, 2, 1)$, $w = (5, -1, -2, 6)$ de $\mathbf{R}^4$.

(a) Mostre que $\langle 3u - 2v, w \rangle = 3\langle u, w \rangle - 2\langle v, w \rangle$.

Por definição,

$$\langle u, w \rangle = 5 - 3 + 8 + 12 = 22 \qquad \text{e} \qquad \langle v, w \rangle = 20 + 2 - 4 + 6 = 24$$

Observe que $3u - 2v = (-5, 13, -16, 4)$. Assim,

$$\langle 3u - 2v, \quad w \rangle = -25 - 13 + 32 + 24 = 18$$

Como era de se esperar, $3\langle u, w \rangle - 2\langle v, w \rangle = 3(22) - 2(24) = 18 = \langle 3u - 2v, \quad w \rangle$.

(b) Normalize u e v.

Por definição,

$$\|u\| = \sqrt{1 + 9 + 16 + 4} = \sqrt{30} \qquad \text{e} \qquad \|v\| = \sqrt{16 + 4 + 4 + 1} = 5$$

Normalizamos u e v e obtemos vetores unitários nas direções e sentidos de u e v, como segue.

$$\hat{u} = \frac{1}{\|u\|} u = \left(\frac{1}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{-4}{\sqrt{30}}, \frac{2}{\sqrt{30}} \right) \qquad \text{e} \qquad \hat{v} = \frac{1}{\|v\|} v = \left(\frac{4}{5}, \frac{-2}{5}, \frac{2}{5}, \frac{1}{5} \right)$$

Espaço funcional *C* [*a*, *b*] e espaço polinomial P(*t*)

A notação $C[a, b]$ é usada para denotar o espaço vetorial de todas as funções contínuas no intervalo fechado $[a, b]$, isto é, com $a \leq t \leq b$. Dadas funções $f(t)$ e $g(t)$ de $C[a, b]$, definimos um produto interno em $C[a, b]$ como segue.

$$\langle f, g \rangle = \int_a^b f(t) g(t) \, dt$$

Esse produto é o *produto interno canônico* de $C[a, b]$.

O espaço vetorial $\mathbf{P}(t)$ de todos os polinômios é um subespaço de $C[a, b]$, para qualquer intervalo $[a, b]$ e, portanto, o produto interno canônico também é um produto interno em $\mathbf{P}(t)$.

Exemplo 7.3

Sejam $f(t) = 3t - 5$ e $g(t) = t^2$ do espaço polinomial $\mathbf{P}(t)$ com produto interno

$$\langle f, g \rangle = \int_0^1 f(t) g(t) \, dt.$$

(a) Encontre $\langle f, g \rangle$.

Temos $f(t)g(t) = 3t^3 - 5t^2$, portanto,

$$\langle f, g \rangle = \int_0^1 (3t^3 - 5t^2) \, dt = \tfrac{3}{4} t^4 - \tfrac{5}{3} t^3 \Big|_0^1 = \tfrac{3}{4} - \tfrac{5}{3} = -\tfrac{11}{12}$$

(b) Encontre $\|f\|$ e $\|g\|$.

Temos $[f(t)]^2 = f(t)f(t) = 9t^2 - 30t + 25$ e $[g(t)]^2 = t^4$. Então

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 (9t^2 - 30t + 25)\, dt = 3t^3 - 15t^2 + 25t\Big|_0^1 = 13$$

$$\|g\|^2 = \langle g, g\rangle = \int_0^1 t^4\, dt = \tfrac{1}{5}t^5\Big|_0^1 = \tfrac{1}{5}$$

Assim, $\|f\| = \sqrt{13}$ e $\|g\| = \sqrt{\frac{1}{5}} = \frac{1}{5}\sqrt{5}$.

Espaço matricial $\mathbf{M} = \mathbf{M}_{m,n}$

Seja $\mathbf{M} = \mathbf{M}_{m,n}$ o espaço vetorial de todas as matrizes reais $m \times n$. Um produto interno de $\mathbf{M}$ é definido por

$$\langle A, B\rangle = \mathrm{tr}(B^T A)$$

onde, como sempre, tr( ) indica o traço, que é a soma dos elementos diagonais. Se $A = [a_{ij}]$ e $B = [b_{ij}]$, então

$$\langle A, B\rangle = \mathrm{tr}(B^T A) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij} \qquad \text{e} \qquad \|A\|^2 = \langle A, A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2$$

Em outras palavras, $\langle A, B\rangle$ é a soma dos produtos das entradas correspondentes de A e B e, em particular, $\langle A, A\rangle$ é a soma dos quadrados das entradas de A.

Espaço de Hilbert

Seja V o espaço vetorial de todas as sequências infinitas $(a_1, a_2, a_3, ...)$ de números reais satisfazendo

$$\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + \cdots < \infty$$

ou seja, a soma infinita converge. A soma e a multiplicação por escalar em V são definidas componente a componente, isto é, se

$$u = (a_1, a_2, \ldots) \qquad \text{e} \qquad v = (b_1, b_2, \ldots)$$

então

$$u + v = (a_1 + b_1,\ a_2 + b_2,\ \ldots) \qquad \text{e} \qquad ru = (ra_1, ra_2, \ldots)$$

Um produto interno de V é definido por

$$\langle u, v\rangle = a_1b_1 + a_2b_2 + \cdots$$

Dado qualquer par de pontos de V, essa soma converge absolutamente. Por isso, esse produto interno está bem definido. Esse espaço com produto interno é denominado *espaço l_2* ou *espaço de Hilbert*.

7.4 DESIGUALDADE DE CAUCHY-SCHWARZ, APLICAÇÕES

A fórmula seguinte (demonstrada no Problema 7.8) é a desigualdade de Cauchy-Schwarz ou desigualdade de Schwarz, utilizada em várias áreas da Matemática.

Teorema 7.1 (Cauchy-Schwarz) Dados vetores u e v quaisquer de um espaço com produto interno V,

$$\langle u, v\rangle^2 \le \langle u, u\rangle\langle v, v\rangle \qquad \text{ou} \qquad |\langle u, v\rangle| \le \|u\|\|v\|$$

Vejamos essa desigualdade em casos específicos.

Exemplo 7.4

(a) Considere quaisquer números reais $a_1, ..., a_n, b_1, ..., b_n$. Pela desigualdade de Cauchy-Schwarz,

$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \le (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

Ou seja, $(u \cdot v)^2 \le \|u\|^2\|v\|^2$, com $u = (a_i)$ e $v = (b_i)$.

(b) Sejam *f* e *g* funções contínuas do intervalo [0, 1]. Então, pela desigualdade de Cauchy-Schwarz,

$$\left[\int_0^1 f(t)g(t)\,dt\right]^2 \leq \int_0^1 f^2(t)\,dt \int_0^1 g^2(t)\,dt$$

Ou seja, $(\langle f, g\rangle)^2 \leq \|f\|^2\|v\|^2$. Aqui, *V* é o espaço com produto interno *C*[0, 1].

O próximo teorema (demonstrado no Problema 7.9) fornece propriedades básicas de uma norma. A demonstração da terceira propriedade requer a desigualdade de Cauchy-Schwarz.

Teorema 7.2 Seja *V* um espaço com produto interno. Então a norma de *V* satisfaz as propriedades listadas a seguir.

[N_1] $\|v\| \geq 0$; e $\|v\| = 0$ se, e só se, $v = 0$.
[N_2] $\|kv\| = |k|\|v\|$.
[N_3] $\|u + v\| \leq \|u\| + \|v\|$.

A propriedade [N_3] é denominada *desigualdade triangular* porque, considerando $u + v$ como o lado do triângulo formado pelos lados *u* e v (conforme Figura 7-1), então [N_3] afirma que o comprimento de um lado de um triângulo não pode ser maior do que a soma dos comprimentos dos outros dois lados.

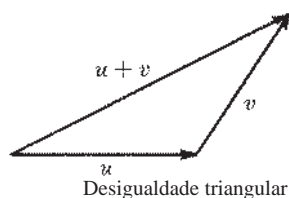



Figura 7-1

Ângulo entre vetores

Dados vetores não nulos *u* e v quaisquer de um espaço com produto interno *V*, definimos o *ângulo entre u* e v como o ângulo θ tal que $0 \leq \theta \leq \pi$ e

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}$$

A desigualdade de Cauchy-Schwarz fornece $-1 \leq \cos\theta \leq 1$, de modo que o ângulo existe e é único.

Exemplo 7.5

(a) Considere os vetores $u = (2, 3, 5)$ e $v = (1, -4, 3)$ em $\mathbf{R}^3$. Então

$$\langle u, v\rangle = 2 - 12 + 15 = 5, \qquad \|u\| = \sqrt{4+9+25} = \sqrt{38}, \qquad \|v\| = \sqrt{1+16+9} = \sqrt{26}$$

e o ângulo θ entre *u* e v é dado por

$$\cos\theta = \frac{5}{\sqrt{38}\sqrt{26}}$$

Observe que θ é um ângulo agudo, pois $\cos\theta$ é positivo.

(b) Sejam $f(t) = 3t - 5$ e $g(t) = t^2$ do espaço polinomial $\mathbf{P}(t)$ com o produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Pelo Exemplo 7.3, temos

$$\langle f, g\rangle = -\tfrac{11}{12}, \qquad \|f\| = \sqrt{13}, \qquad \|g\| = \tfrac{1}{5}\sqrt{5}$$

Então o "ângulo" θ entre *f* e *g* é dado por

$$\cos\theta = \frac{-\frac{11}{12}}{(\sqrt{13})\left(\frac{1}{5}\sqrt{5}\right)} = -\frac{55}{12\sqrt{13}\sqrt{5}}$$

Observe que θ é um ângulo obtuso, pois $\cos\theta$ é negativo.

7.5 ORTOGONALIDADE

Seja V um espaço com produto interno. Dizemos que os vetores $u, v \in V$ são *ortogonais*, ou que u é *ortogonal a* v, se

$$\langle u, v\rangle = 0$$

Essa relação é claramente simétrica, isto é, se u é ortogonal a v, então $\langle v, u\rangle = 0$ e, portanto, v é ortogonal a u. Observamos que $0 \in V$ é ortogonal a cada $v \in V$, pois

$$\langle 0, v\rangle = \langle 0v, v\rangle = 0\langle v, v\rangle = 0$$

Reciprocamente, se u é ortogonal a cada $v \in V$, então $\langle u, u\rangle = 0$ e, portanto, $u = 0$ por [I_3]. Observe que u e v são ortogonais se, e só se, $\cos\theta = 0$, onde θ é o ângulo entre u e v. Ocorre que isso é verdadeiro se, e só se, u e v são "perpendiculares", ou seja, $\theta = \pi/2$ (ou $\theta = 90°$).

Exemplo 7.6

(a) Considere os vetores $u = (1, 1, 1)$, $v = (1, 2, –3)$ e $w = (1, –4, 3)$ de $\mathbf{R}^3$. Então

$$\langle u, v\rangle = 1 + 2 - 3 = 0, \qquad \langle u, w\rangle = 1 - 4 + 3 = 0, \qquad \langle v, w\rangle = 1 - 8 - 9 = -16$$

Assim, u é ortogonal a v e w, mas v e w não são ortogonais.

(b) Considere as funções sen t e cos t no espaço vetorial $C[-\pi, \pi]$ das funções contínuas do intervalo $[-\pi, \pi]$. Então

$$\langle \operatorname{sen} t, \cos t\rangle = \int_{-\pi}^{\pi} \operatorname{sen} t \cos t \, dt = \tfrac{1}{2} \operatorname{sen}^2 t|_{-\pi}^{\pi} = 0 - 0 = 0$$

Assim, sen t e cos t são funções ortogonais do espaço vetorial $C[-\pi, \pi]$.

OBSERVAÇÃO Um vetor $w = (x_1, x_2, \ldots, x_n)$ é ortogonal a $u = (a_1, a_2, \ldots, a_n)$ em $\mathbf{R}^n$ se

$$\langle u, w\rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

Assim, w é ortogonal a u se w satisfaz uma equação homogênea cujos coeficientes são os elementos de u.

Exemplo 7.7 Encontre um vetor não nulo w que seja ortogonal a $u_1 = (1, 2, 1)$ e $u_2 = (2, 5, 4)$ em R^3.

Seja $w = (x, y, z)$. Queremos $\langle u_1, w\rangle = 0$ e $\langle u_2, w\rangle = 0$. Isso fornece o sistema homogêneo

$$\begin{aligned} x + 2y + z &= 0 \\ 2x + 5y + 4z &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y + z &= 0 \\ y + 2z &= 0 \end{aligned}$$

Aqui, z é a única variável livre no sistema escalonado. Tomando $z = 1$, obtemos $y = –2$ e $x = 3$. Assim, $(3, –2, 1)$ é um vetor não nulo que é ortogonal a u_1 e u_2.

Qualquer múltiplo de w também é ortogonal a u_1 e u_2. Normalizando w, obtemos o vetor unitário seguinte que é ortogonal a u_1 e u_2.

$$\hat{w} = \frac{w}{\|w\|} = \left(\frac{3}{\sqrt{14}}, -\frac{2}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right)$$

Complemento ortogonal

Seja S um subconjunto de um espaço com produto interno V. O *complemento ortogonal* de S, denotado por $S^\perp$ (lê-se "esse perpendicular") consiste nos vetores de V que são ortogonais a cada vetor $u \in S$, ou seja,

$$S^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ para cada } u \in S\}$$

Em particular, dado um vetor u particular de V, temos

$$u^{\perp} = \{v \in V : \langle v, u\rangle = 0\}$$

ou seja, $u^{\perp}$ consiste nos vetores de V que são ortogonais ao dado vetor u.

Mostremos que $S^{\perp}$ é um subespaço de V. Claramente, $0 \in S^{\perp}$ porque 0 é ortogonal a cada vetor de V. Suponha, agora, que $v,\ w \in S^{\perp}$. Então, dados escalares a e b quaisquer e um vetor $u \in S$ qualquer, temos

$$\langle av + bw,\ \ u\rangle = a\langle v, u\rangle + b\langle w, u\rangle = a \cdot 0 + b \cdot 0 = 0$$

Assim, $av + bw \in S^{\perp}$ e, portanto, $S^{\perp}$ é um subespaço de V.

Enunciamos formalmente esse resultado.

Proposição 7.3 Seja S um subconjunto de um espaço vetorial V. Então $S^{\perp}$ é um subespaço de V.

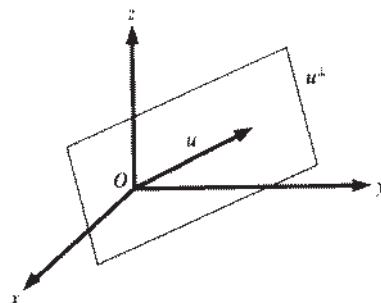
OBSERVAÇÃO 1 Seja u um vetor não nulo de $\mathbf{R}^3$. Então há uma interpretação geométrica de $u^{\perp}$. Mais precisamente, $u^{\perp}$ é o plano de $\mathbf{R}^3$ que passa pela origem O e é perpendicular ao vetor u, conforme Figura 7-2.

Complemento ortogonal $u^{\perp}$

Figura 7-2

OBSERVAÇÃO 2 Seja W o espaço solução de um sistema homogêneo $AX = 0$, em que $A = [a_{ij}]$ é uma matriz $m \times n$ e $X = [x_i]$. Vimos que W é o núcleo da transformação linear $A: \mathbf{R}^n \to \mathbf{R}^m$. Agora podemos dar uma outra interpretação para W, usando a noção de ortogonalidade. Mais precisamente, cada vetor solução $w = (x_1, x_2, ..., x_n)$ é ortogonal a cada linha de A, portanto, W é o complemento ortogonal do espaço linha de A.

Exemplo 7.8 Encontre uma base do subespaço $u^{\perp}$ de $\mathbf{R}^3$, com $u = (1, 3, -4)$.

Observe que $u^{\perp}$ consiste em todos os vetores $w = (x, y, z)$ tais que $\langle u, w\rangle = 0$, ou $x + 3y - 4z = 0$. As variáveis livres são y e z.

(1) Tomando $y = 1$ e $z = 0$, obtemos a solução $w_1 = (-3, 1, 0)$.
(2) Tomando $y = 0$ e $z = 1$, obtemos a solução $w_2 = (4, 0, 1)$.
Os vetores w_1 e w_2 formam uma base do espaço solução da equação e, portanto, uma base de $u^{\perp}$.

Seja W um subespaço de V. Então W e $W^{\perp}$ são, ambos, subespaços de V. O próximo teorema, cuja demonstração (Problema 7.28) requer resultados de seções posteriores, é um resultado fundamental da Álgebra Linear.

Teorema 7.4 Seja W um subespaço de V. Então V é a soma direta de W e $W^{\perp}$, ou seja, $V = W \oplus W^{\perp}$.

7.6 CONJUNTOS ORTOGONAIS E BASES

Considere um conjunto $S = \{u_1, u_2, ..., u_r\}$ de vetores não nulos de um espaço com produto interno V. Dizemos que S é *ortogonal* se dois quaisquer dos vetores de S são ortogonais e dizemos que S é *ortonormal* se S é ortogonal e cada vetor de S é unitário. Isto é,

(i) **Ortogonal** $\langle u_i, u_j\rangle = 0$, com $i \neq j$

(ii) **Ortonormal** $\langle u_i, u_j \rangle = \begin{cases} 0 & \text{com } i \neq j \\ 1 & \text{com } i = j \end{cases}$

Normalizar um conjunto ortogonal S se refere ao processo de multiplicar cada vetor de S pelo recíproco de seu comprimento e, dessa forma, transformando S num conjunto ortonormal de vetores.

Temos os seguintes teoremas.

Teorema 7.5 Se S é um conjunto ortogonal de vetores não nulos, então S é linearmente independente.

Teorema 7.6 (Pitágoras) Seja $\{u_1, u_2, ..., u_r\}$ um conjunto ortogonal de vetores. Então

$$\|u_1 + u_2 + \cdots + u_r\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_r\|^2$$

Esses teoremas são demonstrados nos Problemas 7.15 e 7.16, respectivamente. Aqui, demonstramos o Teorema de Pitágoras no caso especial e conhecido de dois vetores. Mais precisamente, suponha que $\langle u, v \rangle = 0$. Então

$$\|u + v\|^2 = \langle u + v, \ \ u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$$

que dá nosso resultado.

Exemplo 7.9

(a) Seja $E = \{e_1, e_2, e_3\} = \{(1, 0, 0), \ \ (0, 1, 0), \ \ (0, 0, 1)\}$ a base canônica do espaço euclidiano $\mathbf{R}^3$. É claro que

$$\langle e_1, e_2 \rangle = \langle e_1, e_3 \rangle = \langle e_2, e_3 \rangle = 0 \qquad \text{e} \qquad \langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = \langle e_3, e_3 \rangle = 1$$

Assim, E é uma base ortonormal de $\mathbf{R}^3$. Mais geralmente, a base canônica de $\mathbf{R}^n$ é ortonormal, para cada n.

(b) Seja $V = C[-\pi, \pi]$ o espaço vetorial das funções contínuas no intervalo $-\pi \leq t \leq \pi$ com o produto interno definido por $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t)\, dt$. O conjunto a seguir é um exemplo clássico de conjunto ortogonal de V.

$$\{1, \cos t, \cos 2t, \cos 3t, \ldots, \operatorname{sen} t, \operatorname{sen} 2t, \operatorname{sen} 3t, \ldots\}$$

Esse conjunto ortogonal desempenha um papel fundamental na teoria das Séries de Fourier.

Bases ortogonais e combinações lineares, coeficientes de Fourier

Seja S o conjunto constituído pelos três vetores de $\mathbf{R}^3$ dados.

$$u_1 = (1, 2, 1), \qquad u_2 = (2, 1, -4), \qquad u_3 = (3, -2, 1)$$

O leitor pode verificar que esses vetores são ortogonais, portanto, são linearmente independentes. Assim, S é uma base ortogonal de $\mathbf{R}^3$.

Vejamos como escrever $v = (7, 1, 9)$ como uma combinação linear de u_1, u_2, u_3. Primeiro escrevemos v como uma combinação linear de u_1, u_2, u_3 usando incógnitas x_1, x_2, x_3, como segue.

$$v = x_1 u_1 + x_2 u_2 + x_3 u_3 \qquad \text{ou} \qquad (7, 1, 9) = x_1(1, 2, 1) + x_2(2, 1, -4) + x_3(3, -2, 1) \tag{*}$$

Podemos continuar de duas maneiras.

Método 1 Expandimos (*) (como no Capítulo 3) para obter o sistema

$$x_1 + 2x_2 + 3x_3 = 7, \qquad 2x_1 + x_2 - 2x_3 = 1, \qquad x_1 - 4x_2 + x3 = 7$$

Resolvemos o sistema por eliminação gaussiana para obter $x_1 = 3, \ x_2 = -1, \ x_3 = 2$. Assim, $v = 3u_1 - u_2 + 2u_3$.

Método 2 (Este método utiliza a ortogonalidade da base, o que simplifica as contas.) Tomando o produto interno de cada lado de (*) em relação a u_i, obtemos

$$\langle v, u_i\rangle = \langle x_1u_2 + x_2u_2 + x_3u_3,\ \ u_i\rangle \qquad \text{ou} \qquad \langle v, u_i\rangle = x_i\langle u_i, u_i\rangle \qquad \text{ou} \qquad x_i = \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$$

Nesse caso, dois termos foram eliminados, porque u_1, u_2, u_3 é ortogonal. Segue que

$$x_1 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{7+2+9}{1+4+1} = \frac{18}{6} = 3, \qquad x_2 = \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{14+1-36}{4+1+16} = \frac{-21}{21} = -1$$

$$x_3 = \frac{\langle v, u_3\rangle}{\langle u_3, u_3\rangle} = \frac{21-2+9}{9+4+1} = \frac{28}{14} = 2$$

Assim, novamente, obtemos $v = 3u_1 - u_2 + 2u_3$.

O procedimento do Método 2 é válido em geral. A saber, temos o teorema a seguir (demonstrado no Problema 7.17).

Teorema 7.7 Seja $\{u_1, u_2, ..., u_n\}$ uma base ortogonal de V. Então, dado qualquer $v \in V$,

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle}u_n$$

OBSERVAÇÃO O escalar $k_i \equiv \dfrac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$ é denominado *coeficiente de Fourier* de v em relação a u_i, por analogia com um coeficiente da série de Fourier de uma função. Esse escalar também tem uma interpretação geométrica, discutida a seguir.

Projeções

Seja V um espaço com produto interno. Sejam dados um vetor não nulo w de V e um outro vetor v. Queremos encontrar a "projeção de v sobre w" que, conforme indicado na Figura 7-3(a), será um múltiplo cw de w tal que $v' = v - cw$ seja ortogonal a w. Isso significa que

$$\langle v - cw,\ \ w\rangle = 0 \qquad \text{ou} \qquad \langle v, w\rangle - c\langle w, w\rangle = 0 \qquad \text{ou} \qquad c = \frac{\langle v, w\rangle}{\langle w, w\rangle}$$

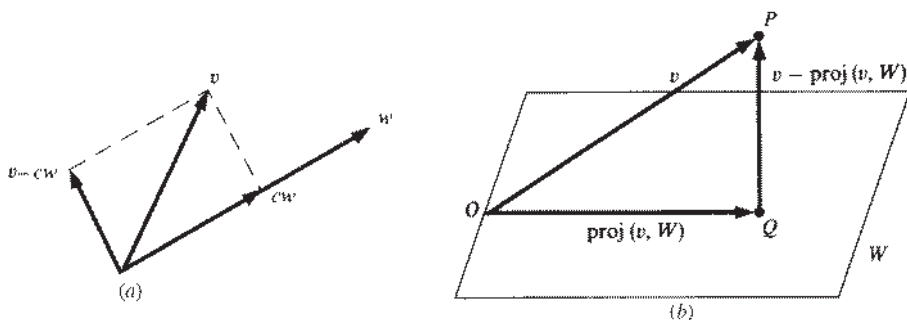



Figura 7-3

Em vista disso, a *projeção de v sobre w* é denotada e definida por

$$\text{proj}(v, w) = cw = \frac{\langle v, w\rangle}{\langle w, w\rangle}w$$

Um tal escalar c é único e é denominado *coeficiente de Fourier* de v em relação a w, ou então, *componente de v ao longo de w*.

Esse conceito é generalizado como segue (ver Problema 7.25).

Teorema 7.8 Sejam $w_1, w_2, ..., w_r$ vetores ortogonais não nulos de V. Dado um vetor v qualquer de V, denotamos

$$v' = v - (c_1 w_1 + c_2 w_2 + \cdots + c_r w_r)$$

com

$$c_1 = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle}, \qquad c_2 = \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle}, \qquad \ldots, \qquad c_r = \frac{\langle v, w_r \rangle}{\langle w_r, w_r \rangle}$$

Então v' é ortogonal a $w_1, w_2, ..., w_r$.

Observe que cada c_i do teorema é o componente (coeficiente de Fourier) de v ao longo do w_i dado.

OBSERVAÇÃO A noção de projeção de um vetor $v \in V$ sobre um subespaço W de V é definida como segue. Pelo Teorema 7.4, $V = W \oplus W^{\perp}$. Logo, v pode ser escrito, de maneira única, na forma

$$v = w + w', \qquad \text{com} \qquad w \in W \qquad \text{e} \qquad w' \in W^{\perp}$$

Definimos w como a *projeção de v sobre W*, denotado por proj(v, W), conforme Figura 7-2(b). Em particular, se W = ger($w_1, w_2, ..., w_r$), em que os w_i formam um conjunto ortogonal, então

$$\text{proj}(v, W) = c_1 w_1 + c_2 w_2 + \cdots + c_r w_r$$

em que cada c_i é o componente de v ao longo de w_i, como antes.

7.7 PROCESSO DE ORTOGONALIZAÇÃO DE GRAM-SCHMIDT

Seja $\{v_1, v_2, ..., v_n\}$ uma base de um espaço com produto interno V. Podemos usar essa base para construir uma base ortogonal $\{w_1, w_2, ..., w_n\}$ de V como segue. Sejam

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2$$

....

$$w_n = v_n - \frac{\langle v_n, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_n, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 - \cdots - \frac{\langle v_n, w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1} \rangle} w_{n-1}$$

Em outras palavras, para $k = 2, 3, ..., n$, definimos

$$w_k = v_k - c_{k1} w_1 - c_{k2} w_2 - \cdots - c_{k,k-1} w_{k-1}$$

onde $c_{ki} = \langle v_k, w_i \rangle / \langle w_i, w_i \rangle$ é o componente de v_k ao longo de w_i. Pelo Teorema 7.8, cada w_k é ortogonal aos w_i que o precedem. Assim, $w_1, w_2, ..., w_n$ formam uma base ortogonal de V, como afirmamos. Normalizando cada w_i, obtemos uma base ortonormal de V.

Essa construção é conhecida como *processo de ortogonalização de Gram-Schmidt*. Convém destacar as observações seguintes.

OBSERVAÇÃO 1 Cada vetor w_k é uma combinação linear de v_k e dos w_i precedentes. Logo, é fácil mostrar, por indução, que cada w_k é uma combinação linear de $v_1, v_2, ..., v_n$.

OBSERVAÇÃO 2 Tomar múltiplos não nulos de vetores não afeta a ortogonalidade, portanto, em contas à mão, pode ser mais simples eliminar as frações em cada novo w_k multiplicando-o por um escalar apropriado antes de passar ao próximo w_{k+1}.

OBSERVAÇÃO 3 Suponha que $u_1, u_2, ..., u_r$ sejam linearmente independentes, de modo que formam uma base de U = ger(u_i). Aplicando o processo de ortogonalização de Gram-Schmidt aos vetores u_i, obtemos uma base ortogonal de U.

Os teoremas seguintes (demonstrados nos Problemas 7.26 e 7.27) utilizam o algoritmo dado e as observações.

Teorema 7.9 Seja $\{v_1, v_2, ..., v_n\}$ uma base qualquer de um espaço com produto interno V. Então existe uma base ortonormal $\{u_1, u_2, ..., u_n\}$ de V tal que a matriz de mudança de base de $\{v_i\}$ para $\{u_i\}$ é triangular, ou seja, para $k = 1, 2, ..., n$,

$$u_k = a_{k1}v_1 + a_{k2}v_2 + \cdots + a_{kk}v_k$$

Teorema 7.10 Seja $S = \{w_1, w_2, ..., w_r\}$ uma base ortogonal de um subespaço W de um espaço vetorial V. Então S pode ser estendida a uma base ortogonal de V, ou seja, podemos encontrar vetores $w_{r+1}, \ldots, w_n$ tais que $\{w_1, w_2, ..., w_n\}$ seja uma base ortogonal de V.

Exemplo 7.10 Aplique o processo de ortogonalização de Gram-Schmidt para encontrar uma base ortogonal e, depois, uma base ortonormal, do subespaço U de $\mathbf{R}^4$ gerado por

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, 2, 4, 5), \qquad v_3 = (1, -3, -4, -2)$$

(1) Denotamos $w_1 = v_1 = (1, 1, 1, 1)$.
(2) Calculamos

$$v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 = v_2 - \frac{12}{4} w_1 = (-2, -1, 1, 2)$$

e denotamos $w_2 = (-2, -1, 1, 2)$.

(3) Calculamos

$$v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 = v_3 - \frac{(-8)}{4} w_1 - \frac{(-7)}{10} w_2 = \left(\tfrac{8}{5}, -\tfrac{17}{10}, -\tfrac{13}{10}, \tfrac{7}{5}\right)$$

Eliminamos as frações para obter $w_3 = (-6, -17, -13, 14)$.

Assim, w_1, w_2, w_3 forma uma base ortogonal de U. Normalizamos esses vetores para obter uma base ortonormal $\{u_1, u_2, u_3\}$ de U. Temos $\|w_1\|^2 = 4$, $\|w_2\|^2 = 10$, $\|w_3\|^2 = 910$, portanto,

$$u_1 = \frac{1}{2}(1, 1, 1, 1), \qquad u_2 = \frac{1}{\sqrt{10}}(-2, -1, 1, 2), \qquad u_3 = \frac{1}{\sqrt{910}}(16, -17, -13, 14)$$

Exemplo 7.11 Seja V o espaço vetorial dos polinômios $f(t)$ com o produto interno $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt$. Aplique o processo de ortogonalização de Gram-Schmidt a $\{1, t, t^2, t^3\}$ para encontrar uma base ortogonal $\{f_0, f_1, f_2, f_3\}$ de $\mathbf{P}_3(t)$ com coeficientes inteiros.

Aqui utilizamos que, para $r + s = n$,

$$\langle t^r, t^s\rangle = \int_{-1}^{1} t^n\,dt = \frac{t^{n+1}}{n+1}\bigg|_{-1}^{1} = \begin{cases} 2/(n+1) & \text{com } n \text{ par} \\ 0 & \text{com } n \text{ ímpar} \end{cases}$$

(1) Denotamos $f_0 = 1$.
(2) Calculamos $t = \dfrac{\langle t, 1\rangle}{\langle 1, 1\rangle}(1) = t - 0 = t$ e denotamos $f_1 = t$.
(3) Calculamos

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^2, t\rangle}{\langle t, t\rangle}(t) = t^2 - \frac{\frac{2}{3}}{2}(1) + 0(t) = t^2 - \tfrac{1}{3}$$

Multiplicamos por 3 para obter $f_2 = 3t^2 = 1$.

(4) Calculamos

$$t^3 - \frac{\langle t^3, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^3, t\rangle}{\langle t, t\rangle}(t) - \frac{\langle t^3,\ 3t^2 - 1\rangle}{\langle 3t^2 - 1,\ 3t^2 - 1\rangle}(3t^2 - 1)$$

$$= t^3 - 0(1) - \frac{\frac{2}{5}}{\frac{2}{3}}(t) - 0(3t^2 - 1) = t^3 - \tfrac{3}{5}t$$

Multiplicamos por 5 para obter $f_3 = 5t^3 - 3t$.

Assim, $\{1,\ \ t,\ \ 3t^2 - 1,\ \ 5t^3 - 3t\}$ é a base ortogonal procurada.

OBSERVAÇÃO Normalizando os polinômios do Exemplo 7.11, com $p(1) = 1$, obtemos os polinômios

$$1,\ t,\ \tfrac{1}{2}(3t^2 - 1),\ \tfrac{1}{2}(5t^3 - 3t)$$

Esses são os quatro primeiros *polinômios de Legendre*, que aparecem no estudo de equações diferenciais.

7.8 MATRIZES ORTOGONAIS E POSITIVAS

Nesta seção discutimos dois tipos de matrizes muito especiais no contexto de espaços com produto interno reais V. Agora, os vetores de $\mathbf{R}^n$ são representados por vetores coluna, com o que o produto interno do espaço euclidiano $\langle u, v\rangle = u^T v$ passa ser escrito como $\mathbf{R}^n$.

Matrizes ortogonais

Uma matriz real P é denominada *ortogonal* se P é não singular e $P^{-1} = P^T$ ou, em outras palavras, se $PP^T = P^T P = I$. Em primeiro lugar, lembremos (Teorema 2.6) uma caracterização importante dessas matrizes.

Teorema 7.11 Seja P uma matriz real. São equivalentes as afirmações seguintes. (a) P é ortogonal; (b) as linhas de P formam um conjunto ortonormal; (c) as colunas de P formam um conjunto ortonormal.

(Esse teorema só é válido com o produto interno canônico de $\mathbf{R}^n$, não sendo verdadeiro usando algum outro produto interno de $\mathbf{R}^n$.)

Exemplo 7.12

(a) Seja $P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}$. As linhas de P são ortogonais umas às outras e todas são vetores unitários. Assim, P é uma matriz ortogonal.

(b) Seja P uma matriz ortogonal 2×2. Então, para algum número real θ, temos

$$P = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \qquad \text{ou} \qquad P = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

Os dois teoremas a seguir (demonstrados nos Problemas 7.37 e 7.38) mostram relações importantes entre as matrizes ortogonais e as bases ortonormais de um espaço com produto interno real V.

Teorema 7.12 Sejam $E = \{e_i\}$ e $E = \{e_i\}$ e $E' = \{e'_i\}$ bases ortonormais de V e P a matriz de mudança de base da base E para a base E'. Então P é ortogonal.

Teorema 7.13 Sejam $\{e_1, ..., e_n\}$ uma base ortonormal de um espaço com produto interno V e $P = [a_{ij}]$ uma matriz ortogonal. Então os n vetores seguintes formam uma base ortonormal de V.

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, 2, \ldots, n$$

Matrizes positivas

Seja A uma matriz real simétrica, isto é, $A^T = A$. Dizemos que A é *positiva* se, para cada vetor não nulo u de $\mathbf{R}^n$,

$$\langle u, Au\rangle = u^T Au > 0$$

No Capítulo 12 veremos algoritmos que nos permitem decidir se uma matriz dada é positiva, ou não. Contudo, para matrizes 2×2, temos um critério simples que enunciamos no teorema seguinte (demonstrado no Problema 7.43).

Teorema 7.14 Uma matriz real 2×2 simétrica $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ é positiva se, e só se, os elementos diagonais a e d são positivos e seu determinante $|A| = ad - bc = ad - b^2$ é positivo.

Exemplo 7.13 Considere as matrizes simétricas seguintes.

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -2 \\ -2 & -3 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$$

A não é positiva, pois $|A| = 4 - 9 = -5$ é negativo. B não é positiva, pois a entrada diagonal –3 é negativa. Contudo, C é positiva, pois seus elementos diagonais 1 e 5 são positivos e seu determinante $|C| = 5 - 4 = 1$ também é positivo.

Vale o teorema seguinte (demonstrado no Problema 7.44).

Teorema 7.15 Seja A uma matriz real positiva. Então a função $\langle u, v\rangle = u^T Av$ é um produto interno de $\mathbf{R}^n$.

Representação matricial de um produto interno (opcional)

O Teorema 7.15 afirma que toda matriz real positiva determina um produto interno de $\mathbf{R}^n$. Nesta seção fornecemos a recíproca desse resultado.

Seja V um espaço com produto interno real com base $S = \{u_1, u_2, ..., u_n\}$. A matriz

$$A = [a_{ij}], \qquad \text{com} \qquad a_{ij} = \langle u_i, u_j\rangle$$

é denominada *representação matricial do produto interno de V em relação à base S*.

Observe que A é simétrica, porque o produto interno é simétrico, isto é, $\langle u_i, u_j\rangle = \langle u_j, u_i\rangle$. A também depende do produto interno de V e da base S de V. Além disso, se S é uma base ortogonal, então A é diagonal e, se S é uma base ortonormal, então A é a matriz identidade.

Exemplo 7.14 Os vetores $u_1 = (1, 1, 0)$, $u_2 = (1, 2, 3)$, $u_3 = (1, 3, 5)$ formam uma base S do espaço euclidiano $\mathbf{R}^3$. Encontre a matriz A que representa o produto interno de $\mathbf{R}^3$ em relação à base S.

Inicialmente calculamos cada $\langle u_i, u_j\rangle$, obtendo

$$\begin{array}{lll} \langle u_1, u_1\rangle = 1 + 1 + 0 = 2, & \langle u_1, u_2\rangle = 1 + 2 + 0 = 3, & \langle u_1, u_3\rangle = 1 + 3 + 0 = 4 \\ \langle u_2, u_2\rangle = 1 + 4 + 9 = 14, & \langle u_2, u_3\rangle = 1 + 6 + 15 = 22, & \langle u_3, u_3\rangle = 1 + 9 + 25 = 35 \end{array}$$

Então $A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 14 & 22 \\ 4 & 22 & 35 \end{bmatrix}$. Como era de se esperar, A é simétrica.

Valem os teoremas seguintes (demonstrados nos Problemas 7.45 e 7.46, respectivamente).

Teorema 7.16 Seja A a representação matricial de um produto interno em relação a uma base S de V. Então, dados quaisquer vetores $u, v \in V$, temos

$$\langle u, v\rangle = [u]^T A[v]$$

onde $[u]$ e $[v]$ denotam os vetores (coluna) de coordenadas em relação à base S.

Teorema 7.17 Seja A a representação matricial de algum produto interno de V. Então A é uma matriz positiva.

7.9 ESPAÇOS COMPLEXOS COM PRODUTO INTERNO

Nesta seção consideramos espaços vetoriais sobre o corpo complexo $\mathbf{C}$. Inicialmente, relembramos algumas propriedades dos números complexos (Seção 1.7), especialmente as relações entre um número complexo $\bar{z} = a + bi$, com $a, b \in \mathbf{R}$, e seu conjugado complexo $\bar{z} = a - bi$, como segue.

$$z\bar{z} = a^2 + b^2, \qquad |z| = \sqrt{a^2 + b^2}, \qquad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \qquad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \qquad \bar{\bar{z}} = z$$

Além disso, z é real se, e só se, $\bar{z} = z$.

Temos a definição seguinte.

DEFINIÇÃO Seja V um espaço vetorial sobre $\mathbf{C}$. Suponha que a cada par de vetores $u, v \in V$ esteja associado um número complexo, denotado por $\langle u, v\rangle$. Essa função é denominada *produto interno (complexo)* de V se satisfizer os axiomas seguintes.

$[I_1^*]$ **(Linearidade)** $\langle au_1 + bu_2,\ v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$
$[I_2^*]$ **(Simetria Conjugada)** $\langle u, v\rangle = \overline{\langle v, u\rangle}$
$[I_3^*]$ **(Positividade)** $\langle u, u\rangle \geq 0$ e $\langle u, u\rangle = 0$ se, e só se, $u = 0$.

O espaço vetorial V sobre $\mathbf{C}$ com um produto interno é denominado *espaço com produto interno (complexo)*.

Observe que única distinção entre um produto interno real e um complexo é o segundo axioma $[I_2^*]$.

O axioma $[I_1^*]$ da linearidade é equivalente às duas condições

$$(a) \quad \langle u_1 + u_2,\ v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle, \qquad (b) \quad \langle ru, v\rangle = r\langle u, v\rangle$$

Por outro lado, aplicando $[I_1^*]$ e $[I_2^*]$, obtemos

$$\langle u, rv\rangle = \overline{\langle rv, u\rangle} = \overline{r\langle v, u\rangle} = \bar{r}\,\overline{\langle v, u\rangle} = \bar{r}\langle u, v\rangle$$

Assim, devemos tomar o conjugado de um número complexo que for retirado da segunda posição de um produto interno complexo. De fato (Problema 7.47), o produto interno é *conjugado linear* na segunda posição, isto é,

$$\langle u,\ av_1 + bv_2\rangle = \bar{a}\langle u, v_1\rangle + \bar{b}\langle u, v_2\rangle$$

Combinando a linearidade na primeira posição com a linearidade conjugada da segunda, obtemos, por indução,

$$\left\langle \sum_i a_i u_i,\ \sum_j b_j v_j \right\rangle = \sum_{i,j} a_i \overline{b_j} \langle u_i, v_j\rangle$$

Convém destacar as observações seguintes.

OBSERVAÇÃO 1 O axioma $[I_1^*]$ sozinho implica $\langle 0, 0\rangle = \langle 0v, 0\rangle = 0\langle v, 0\rangle = 0$. Assim, $[I_1^*]$, $[I_2^*]$ e $[I_3^*]$ são equivalentes a $[I_1^*]$ e $[I_2^*]$ juntos com o axioma seguinte.

$[I_3^{*\prime}]$ Se $u \neq 0$, então $\langle u, u\rangle > 0$.

Ou seja, uma função que satisfaz $[I_1^*]$, $[I_2^*]$ e $[I_3^{*\prime}]$ é um produto interno (complexo) de V.

OBSERVAÇÃO 2 Por $[I_2^*]$, $\langle u, u\rangle = \overline{\langle u, u\rangle}$. Assim, $\langle u, u\rangle$ sempre é real. Por $[I_3^*]$, $\langle u, u\rangle$ sempre é não negativo e, portanto, existe a raiz quadrada não negativa desse valor. Como no caso de espaços com produto interno real, definimos $\|u\| = \sqrt{\langle u, u\rangle}$ como a *norma* ou *comprimento* de u.

OBSERVAÇÃO 3 Além da norma, definimos, como antes, os conceitos de ortogonalidade, complemento ortogonal e conjuntos ortogonais e ortonormais. Na verdade, as definições de distância, coeficientes de Fourier e projeções são idênticas às do caso real.

Exemplo 7.15 (Espaço euclidiano complexo $\mathbf{C}^n$). Dado $V = \mathbf{C}^n$, sejam $u = (z_i)$ e $v = (w_i)$ vetores de $\mathbf{C}^n$. Então

$$\langle u, v\rangle = \sum_k z_k \overline{w_k} = z_1\overline{w_1} + z_2\overline{w_2} + \cdots + z_n\overline{w_n}$$

é um produto interno de V, denominado *produto interno canônico* de $\mathbf{C}^n$. Com esse produto interno, dizemos que V é o espaço euclidiano complexo. Salvo menção explícita em contrário, vamos sempre supor definido esse produto interno em $\mathbf{C}^n$. Considerando u e v como vetores coluna, esse produto interno pode
ser definido por

$$\langle u, v\rangle = u^T\bar{v}$$

onde, como com matrizes, $\bar{v}$ significa o conjugado de cada entrada de v. Se u e v são reais, temos $\overline{w_i} = w_i$ e, nesse caso, o produto interno se reduz ao análogo de $\mathbf{R}^n$.

Exemplo 7.16

(a) Seja V o espaço vetorial das funções contínuas complexas definidas no intervalo (real) $a \leq t \leq b$. O *produto interno canônico* de V é dado por

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt$$

(b) Seja U o espaço vetorial das matrizes $m \times n$ sobre $\mathbf{C}$. Sejam $A = (z_{ij})$ e $B = (w_{ij})$ elementos de U. O *produto interno canônico* de U é dado por

$$\langle A, B\rangle = \mathrm{tr}(B^H A) = \sum_{i=1}^{m}\sum_{j=1}^{n} \bar{w}_{ij} z_{ij}$$

Como antes, $B^H = \bar{B}^T$, ou seja, B^H é a matriz conjugada transposta de B.

A seguir, apresentamos uma lista de teoremas relativos a espaços com produto interno complexos que são análogos àqueles do caso real. Aqui, as matrizes hermitianas A (isto é, matrizes tais que $A^H = \bar{A}^T = A$), desempenham o mesmo papel que as matrizes simétricas A (isto é, matrizes tais que $A^T = A$) no caso real. (O Teorema 7.18 é demonstrado no Problema 7.50.)

Teorema 7.18 (Cauchy-Schwarz) Seja V um espaço com produto interno complexo. Então

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

Teorema 7.19 Seja W um subespaço de um espaço com produto interno complexo V. Então $V = W \oplus W^{\perp}$.

Teorema 7.20 Suponha que $\{u_1, u_2, ..., u_n\}$ seja uma base de um espaço com produto interno complexo V. Então, dado qualquer $v \in V$,

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle}u_n$$

Teorema 7.21 Suponha que $\{u_1, u_2, ..., u_n\}$ seja uma base de um espaço com produto interno complexo V. Seja $A = [a_{ij}]$ a matriz complexa definida por $a_{ij} = \langle u_i, u_j\rangle$. Então, dados quaisquer $u, v \in V$,

$$\langle u, v\rangle = [u]^T A\overline{[v]}$$

onde $[u]$ e $[v]$ denotam os vetores coluna de coordenadas na base $\{u_i\}$. (*Observação*: dizemos que essa matriz A representa o produto interno de V.)

Teorema 7.22 Seja A uma matriz hermitiana (isto é, $A^H = \bar{A}^T = A$) tal que $X^TA\bar{X}$ seja real e positivo, para qualquer vetor não nulo $X \in \mathbf{C}^n$. Então, $\langle u, v\rangle = u^TA\bar{v}$ é um produto interno de $\mathbf{C}^n$.

Teorema 7.23 Seja A a matriz que representa um produto interno de V. Então A é hermitiana e X^TAX é real e positivo, para qualquer vetor não nulo X de $\mathbf{C}^n$.

7.10 ESPAÇOS VETORIAIS NORMADOS (OPCIONAL)

Começamos com uma definição.

DEFINIÇÃO Seja V um espaço vetorial real ou complexo. Suponha que a cada $v \in V$ esteja associado um número real, denotado por $\|v\|$. Essa função $\|\cdot\|$ é denominada *norma* de V se satisfizer os axiomas seguintes.

[N_1] $\|v\| \geq 0$ e $\|v\| = 0$ se, e só se, $v = 0$.
[N_2] $\|kv\| = |k|\|v\|$.
[N_3] $\|u + v\| \leq \|u\| + \|v\|$.

Um espaço vetorial V com uma norma é denominado *espaço vetorial normado*.

Seja V um espaço vetorial normado. A *distância* entre dois vetores u e v de V é denotada e definida por

$$d(u, v) = \|u - v\|$$

O teorema seguinte (demonstrado no Problema 7.56) é a razão principal pela qual dizemos que $d(u, v)$ é a distância entre u e v.

Teorema 7.24 Seja V um espaço vetorial normado. Então a função $d(u, v) = \|u - v\|$ satisfaz os três axiomas de espaço métrico, como segue.

[M_1] $d(u, v) \geq 0$ e $d(u, v) = 0$ se, e só se, $u = v$.
[M_2] $d(u, v) = d(v, u)$.
[M_3] $d(u, v) \leq d(u, w) + d(w, v)$.

Espaços vetoriais normados e espaços com produto interno

Seja V é um espaço com produto interno. A norma de um vetor v de V foi definida por

$$\|v\| = \sqrt{\langle v, v\rangle}$$

Prova-se (Teorema 7.2) que essa norma satisfaz [N_1], [N_2] e [N_3]. Assim, todo espaço com produto interno V é um espaço normado. Por outro lado, podem existir normas num espaço vetorial V que não tenham sua origem num produto interno de V.

Normas de $\mathbf{R}^n$ e $\mathbf{C}^n$

Três normas importantes de $\mathbf{R}^n$ e $\mathbf{C}^n$ são definidas como segue.

$$\|(a_1, \ldots, a_n)\|_\infty = \max(|a_i|)$$
$$\|(a_1, \ldots, a_n)\|_1 = |a_1| + |a_2| + \cdots + |a_n|$$
$$\|(a_1, \ldots, a_n)\|_2 = \sqrt{|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2}$$

(Observe que usamos índices para diferenciar entre essas três normas.) As normas $\|\cdot\|_\infty$, $\|\cdot\|_1$ e $\|\cdot\|_2$ são denominadas, respectivamente, *norma infinito, norma 1* e *norma 2*. Observe que $\|\cdot\|_2$ é a norma de $\mathbf{R}^n$ (respectivamente, $\mathbf{C}^n$) induzida pelo produto interno canônico de $\mathbf{R}^n$ (respectivamente, $\mathbf{C}^n$). Vamos denotar as respectivas funções distância por d_∞, d_1, d_2.

Exemplo 7.17 Considere os vetores $u = (1, –5, 3)$ e $v = (4, 2, –3)$ de $\mathbf{R}^3$.

(a) A norma infinito escolhe o máximo dentre os valores absolutos das entradas. Logo,

$$\|u\|_\infty = 5 \qquad \text{e} \qquad \|v\|_\infty = 4$$

(b) A norma 1 soma os valores absolutos das entradas. Logo,

$$\|u\|_1 = 1 + 5 + 3 = 9 \qquad \text{e} \qquad \|v\|_1 = 4 + 2 + 3 = 9$$

(c) A norma 2 é igual à raiz quadrada da soma dos quadrados das entradas (isto é, a norma induzida pelo produto interno canônico de $\mathbf{R}^3$). Assim,

$$\|u\|_2 = \sqrt{1+25+9} = \sqrt{35} \qquad \text{e} \qquad \|v\|_2 = \sqrt{16+4+9} = \sqrt{29}$$

(d) Como $u - v = (1 - 4, -5 - 2, 3 + 3) = (-3, -7, 6)$, temos

$$d_\infty(u, v) = 7, \qquad d_1(u, v) = 3 + 7 + 6 = 16, \qquad d_2(u, v) = \sqrt{9 + 49 + 36} = \sqrt{94}$$

Exemplo 7.18 Considere o plano cartesiano $\mathbf{R}^2$ mostrado na Figura 7-4.

(a) Seja D_1 o conjunto dos pontos $u = (x, y)$ de $\mathbf{R}^2$ tais que $\|u\|_2 = 1$. Então D_1 consiste nos pontos (x, y) tais que $\|u\|_2^2 = x^2 + y^2 = 1$. Assim, D_1 é o círculo unitário, conforme Figura 7-4.

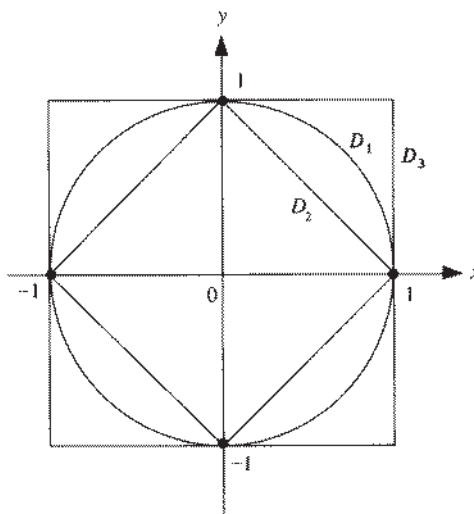



Figura 7-4

(b) Seja D_2 o conjunto dos pontos $u = (x, y)$ de $\mathbf{R}^2$ tais que $\|u\|_1 = 1$. Então D_2 consiste nos pontos (x, y) tais que $\|u\|_1 = |x| + |y| = 1$. Assim, D_2 é um quadrado girado inscrito no círculo unitário, conforme Figura 7-4.

(c) Seja D_3 o conjunto dos pontos $u = (x, y)$ de $\mathbf{R}^2$ tais que $\|u\|_\infty = 1$. Então D_3 consiste nos pontos (x, y) tais que $\|u\|_\infty = \max(|x|, |y|) = 1$. Assim, D_3 é um quadrado horizontal que circunscreve o círculo unitário, conforme Figura 7-4.

Normas de *C*[*a*, *b*]

Considere o espaço vetorial $V = C[a, b]$ das funções reais contínuas do intervalo $a \le t \le b$. Vimos que um produto interno de V é dado por

$$\langle f, g \rangle = \int_a^b f(t)g(t)\, dt$$

Assim, esse produto interno define uma norma de $V = C[a, b]$ (que é análoga à norma $\|\cdot\|_2$ de $\mathbf{R}^n$), como segue.

$$\|f\|_2 = \sqrt{\int_a^b [f(t)]^2\, dt}$$

Outras normas de $V = C[a, b]$ são

$$\|f\|_1 = \int_a^b |f(t)|\, dt \qquad \text{e} \qquad \|f\|_\infty = \max(|f(t)|)$$

As interpretações geométricas dessas duas normas e suas funções distância correspondentes são dadas a seguir.

A primeira norma é mostrada na Figura 7-5. Aqui,

$\|f\|_1 =$ área entre a função $|f|$ e o eixo t

$d_1(f, g) =$ área entre as funções f e g

(a) $\|f\|_1$ é a área destacada

(b) $d_1(f, g)$ é a área destacada

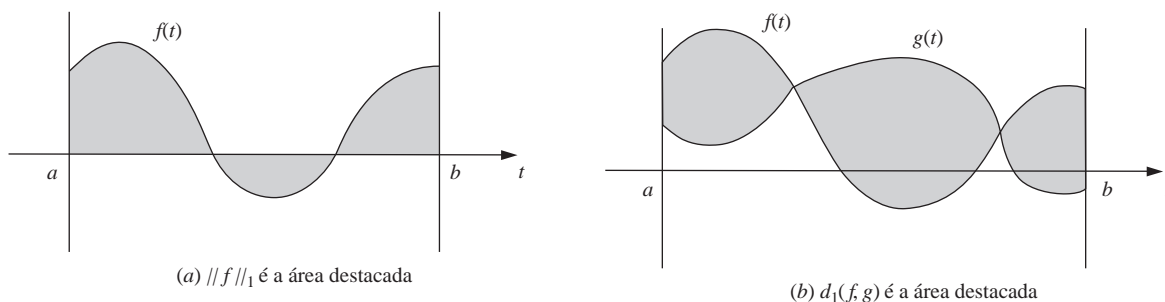



Figura 7-5

Essa norma é análoga à norma $\|\cdot\|_1$ de $\mathbf{R}^n$.

A segunda norma é mostrada na Figura 7-6. Aqui,

$\|f\|_\infty =$ distância máxima entre f e o eixo t
$d_\infty(f, g) =$ distância máxima entre f e g

Essa norma é análoga à norma $\|\cdot\|_\infty$ de $\mathbf{R}^n$.

(a) $\|f\|_\infty$

(b) $d_\infty(f, g)$

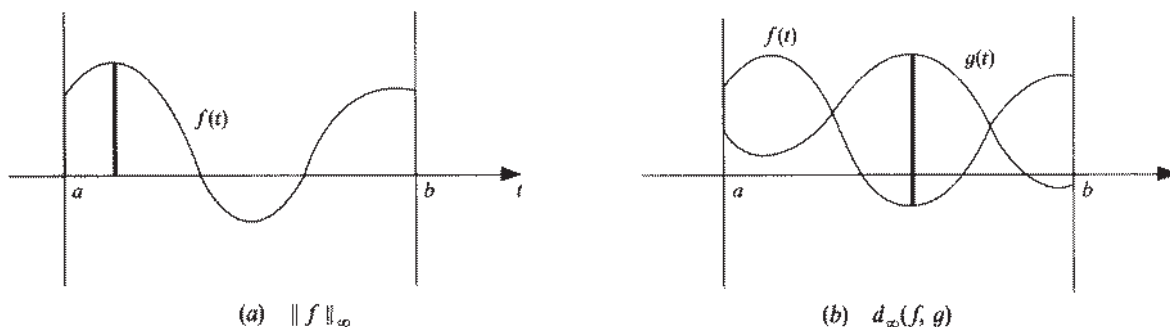



Figura 7-6

Problemas Resolvidos

Produtos internos

7.1 Expanda as expressões dadas.

(a) $\langle 5u_1 + 8u_2,\ \ 6v_1 - 7v_2 \rangle$,
(b) $\langle 3u + 5v,\ \ 4u - 6v \rangle$,
(c) $\|2u - 3v\|^2$

Use linearidade em ambas posições e, quando possível, a simetria $\langle u, v \rangle = \langle v, u \rangle$.

(a) Tomamos o produto interno de cada termo à esquerda com cada termo à direita.

$$\begin{aligned}\langle 5u_1 + 8u_2,\ \ 6v_1 - 7v_2 \rangle &= \langle 5u_1, 6v_1 \rangle + \langle 5u_1, -7v_2 \rangle + \langle 8u_2, 6v_1 \rangle + \langle 8u_2, -7v_2 \rangle \\ &= 30\langle u_1, v_1 \rangle - 35\langle u_1, v_2 \rangle + 48\langle u_2, v_1 \rangle - 56\langle u_2, v_2 \rangle\end{aligned}$$

[*Observação*: confira a semelhança entre essa expansão e a de $(5a–8b)(6c–7d)$ na Álgebra elementar.]

(b) $\langle 3u+5v,\ 4u-6v\rangle = 12\langle u,u\rangle - 18\langle u,v\rangle + 20\langle v,u\rangle - 30\langle v,v\rangle$
$= 12\langle u,u\rangle + 2\langle u,v\rangle - 30\langle v,v\rangle$

(c) $\|2u-3v\|^2 = \langle 2u-3v,\ 2u-3v\rangle = 4\langle u,u\rangle - 6\langle u,v\rangle - 6\langle v,u\rangle + 9\langle v,v\rangle$
$= 4\|u\|^2 - 12(u,v) + 9\|v\|^2$

7.2 Considere os vetores $u=(1,2,4),\ v=(2,-3,5),\ w=(4,2,-3)$ de $\mathbf{R}^3$. Encontre

(a) $u\cdot v$, (b) $u\cdot w$, (c) $v\cdot w$, (d) $(u+v)\cdot w$, (e) $\|u\|$, (f) $\|v\|$.

(a) Multiplicamos componentes correspondentes e somamos para obter $u\cdot v = 2-6+20=16$.

(b) $u\cdot w = 4+4-12=-4$.

(c) $v\cdot w = 8-6-15=-13$.

(d) Calculamos $u+v=(3,-1,9)$ e, então, $(u+v)\cdot w = 12-2-27=-17$. Alternativamente, usando $[I_1]$, $(u+v)\cdot w = u\cdot w + v\cdot w = -4-13=-17$.

(e) Primeiro calculamos $\|u\|^2$ elevando os componentes de u ao quadrado e somando, como segue.

$$\|u\|^2 = 1^2+2^2+4^2 = 1+4+16=21, \qquad \text{e, portanto,} \qquad \|u\| = \sqrt{21}$$

(f) $\|v\|^2 = 4+9+25=38$ e, portanto, $\|v\|=\sqrt{38}$.

7.3 Verifique que define um produto interno de $\mathbf{R}^2$.

$$\langle u,v\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2, \qquad \text{com} \qquad u=(x_1,x_2),\quad v=(y_1,y_2)$$

Argumentamos com matrizes. Podemos escrever $\langle u,v\rangle$ em notação matricial como segue.

$$\langle u,v\rangle = u^TAv = [x_1,x_2]\begin{bmatrix} 1 & -1 \\ -1 & 3\end{bmatrix}\begin{bmatrix} y_1 \\ y_2\end{bmatrix}$$

Como A é real e simétrica, basta mostrar que A é positiva. Os elementos 1 e 3 da diagonal são positivos e o determinante $\|A\| = 3-1=2$ é positivo. Assim, pelo Teorema 7.14, A é positiva. Por isso, o Teorema 7.15 garante que $\langle u,v\rangle$ é um produto interno.

7.4 Considere os vetores $u=(1,5)$ e $v=(3,4)$. Encontre

(a) $\langle u,v\rangle$ em relação ao produto interno canônico de $\mathbf{R}^2$.

(b) $\langle u,v\rangle$ em relação ao produto interno de $\mathbf{R}^2$ do Problema 7.3.

(c) $\|v\|$ usando o produto interno canônico de $\mathbf{R}^2$.

(d) $\|v\|$ usando o produto interno de $\mathbf{R}^2$ do Problema 7.3.

(a) $\langle u,v\rangle = 3+20=23$.

(b) $\langle u,v\rangle = 1\cdot 3 - 1\cdot 4 - 5\cdot 3 + 3\cdot 5\cdot 4 = 3-4-15+60=44$.

(c) $\|v\|^2 = \langle v,v\rangle = \langle(3,4),(3,4)\rangle = 9+16=25$ e, portanto, $\|v\|=5$.

(d) $\|v\|^2 = \langle v,v\rangle = \langle(3,4),(3,4)\rangle = 9-12-12+48=33$ e, portanto, $\|v\|=\sqrt{33}$.

7.5 Considere os polinômios de $\mathbf{P}(t)$ a seguir com o produto interno $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$.

$$f(t)=t+2, \qquad g(t)=3t-2, \qquad h(t)=t^2-2t-3$$

(a) Encontre $\langle f,g\rangle$ e $\langle f,h\rangle$.

(b) Encontre $\|f\|$ e $\|g\|$.

(c) Normalize f e g.

(a) Integramos como segue.

$$\langle f, g\rangle = \int_0^1 (t+2)(3t-2)\ dt = \int_0^1 (3t^2+4t-4)\ dt = \left(t^3+2t^2-4t\right)\Big|_0^1 = -1$$

$$\langle f, h\rangle = \int_0^1 (t+2)(t^2-2t-3)\ dt = \left(\frac{t^4}{4}-\frac{7t^2}{2}-6t\right)\Big|_0^1 = -\frac{37}{4}$$

(b) $\langle f, f\rangle = \int_0^1 (t+2)(t+2)\ dt = \frac{19}{3}$; logo, $\|f\| = \sqrt{\frac{19}{3}} = \frac{1}{3}\sqrt{57}$.

$$\langle g, g\rangle = \int_0^1 (3t-2)(3t-2) = 1; \qquad \text{logo,} \qquad \|g\| = \sqrt{1} = 1$$

(c) Como $\|f\| = \frac{1}{3}\sqrt{57}$ e g já é um vetor unitário, temos

$$\hat{f} = \frac{1}{\|f\|}f = \frac{3}{\sqrt{57}}(t+2) \qquad \text{e} \qquad \hat{g} = g = 3t-2$$

7.6 Encontre θ, sendo θ o ângulo entre

(a) $u = (1, 3, -5, 4)$ e $v = (2, -3, 4, 1)$ em $\mathbf{R}^4$.

(b) $A = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, com $\langle A, B\rangle = \text{tr}(B^TA)$

Use $\cos\theta = \dfrac{\langle u, v\rangle}{\|u\|\|v\|}$

(a) Calculamos

$$\langle u, v\rangle = 2-9-20+4 = -23, \qquad \|u\|^2 = 1+9+25+16 = 51, \qquad \|v\|^2 = 4+9+16+1 = 30$$

Assim, $$\cos\theta = \frac{-23}{\sqrt{51}\sqrt{30}} = \frac{-23}{3\sqrt{170}}$$

(b) Usamos $\langle A, B\rangle = \text{tr}(B^TA) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}b_{ij}$, a soma dos produtos das entradas correspondentes.

$$\langle A, B\rangle = 9+16+21+24+25+24 = 119$$

Usamos $\|A\|^2 = \langle A, A\rangle = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$, a soma dos quadrados de todas as entradas de A.

$$\|A\|^2 = \langle A, A\rangle = 9^2+8^2+7^2+6^2+5^2+4^2 = 271 \qquad \text{e, portanto,} \qquad \|A\| = \sqrt{271}$$

$$\|B\|^2 = \langle B, B\rangle = 1^2+2^2+3^2+4^2+5^2+6^2 = 91 \qquad \text{e, portanto,} \qquad \|B\| = \sqrt{91}$$

Assim, $$\cos\theta = \frac{119}{\sqrt{271}\sqrt{91}}$$

7.7 Verifique as igualdades seguintes.

(a) Lei do paralelogramo (Figura 7-7): $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

(b) Forma polar de $\langle u, v\rangle$ (mostrando que o produto interno pode ser obtido a partir da função norma):

$$\langle u, v\rangle = \tfrac{1}{4}(\|u+v\|^2 - \|u-v\|^2).$$

Expandimos como segue.

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \tag{1}$$
$$\|u-v\|^2 = \langle u-v, u-v\rangle = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2 \tag{2}$$

Somando (1) com (2), obtemos a lei do paralelogramo (a). Subtraindo (2) de (1), obtemos

$$\|u+v\|^2 - \|u-v\|^2 = 4\langle u, v\rangle$$

Dividindo por 4, obtemos a forma polar (real) (b).

7.8 Demonstre o Teorema 7.1 (Cauchy-Schwarz). Dados vetores u e v quaisquer de um espaço com produto interno real V, $\langle u, u\rangle^2 \leq \langle u, u\rangle\langle v, v\rangle$ ou $|\langle u, v\rangle| \leq \|u\|\|v\|$.

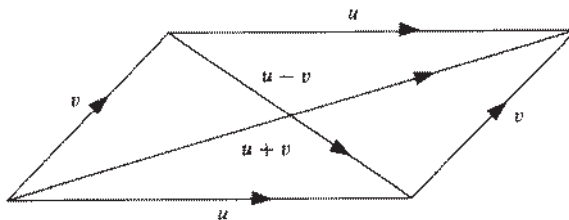



Figura 7-7

Dado um número real t,

$$\langle tu+v,\ \ tu+v\rangle = t^2\langle u,u\rangle + 2t\langle u,v\rangle + \langle v,v\rangle = t^2\|u\|^2 + 2t\langle u,v\rangle + \|v\|^2$$

Sejam $a=\|u\|^2$, $b=2\langle u,v\rangle$, $c=\|v\|^2$. Como $\|tu+v\|^2 \geq 0$, temos

$$at^2+bt+c\geq 0$$

para cada valor de t. Isso significa que o polinômio quadrático não pode ter duas raízes reais, o que implica que $b^2-4ac\leq 0$ ou $b^2\leq 4ac$. Assim,

$$4\langle u,v\rangle^2 \leq 4\|u\|^2\|v\|^2$$

Dividindo por 4, obtemos nosso resultado.

7.9 Demonstre o Teorema 7.2. A norma de um espaço com produto interno V satisfaz

(a) $[N_1]$ $\|v\|\geq 0$ e $\|v\|=0$ se, e só se, $v=0$.

(b) $[N_2]$ $\|kv\|=|k|\|v\|$.

(c) $[N_3]$ $\|u+v\|\leq\|u\|+\|v\|$.

(a) Se $v\neq 0$, então $\langle v,v\rangle>0$ e, portanto, $\|v\|=\sqrt{\langle v,v\rangle}>0$. Se $v=0$, então $\langle 0,0\rangle=0$. Consequentemente, $\|0\|=\sqrt{0}=0$. Assim, $[N_1]$ é válido.

(b) Temos $\|kv\|^2=\langle kv,kv\rangle=k^2\langle v,v\rangle=k^2\|v\|^2$. Tomando a raiz quadrada de ambos lados, resulta $[N_2]$.

(c) Usando a desigualdade de Cauchy-Schwarz, obtemos

$$\begin{aligned}\|u+v\|^2 &= \langle u+v,\ \ u+v\rangle = \langle u,u\rangle+\langle u,v\rangle+\langle u,v\rangle+\langle v,v\rangle\\ &\leq \|u\|^2+2\|u\|\|v\|+\|v\|^2 = (\|u\|+\|v\|)^2\end{aligned}$$

Tomando a raiz quadrada de ambos lados, resulta $[N_3]$.

Ortogonalidade, complementos ortogonais, conjuntos ortogonais

7.10 Encontre k tal que os vetores $u=(1, 2, k, 3)$ e $v=(3, k, 7, -5)$ de $\mathbf{R}^4$ sejam ortogonais.

Calculamos

$$\langle u,v\rangle = (1,2,k,3)\cdot(3,k,7,-5) = 3+2k+7k-15 = 9k-12$$

Tomando $\langle u,v\rangle = 9k-12=0$, obtemos $k=\frac{4}{3}$.

7.11 Seja W o subespaço de $\mathbf{R}^5$ gerado por $u=(1, 2, 3, -1, 2)$ e $v=(2, 4, 7, 2, -1)$. Encontre uma base do complemento ortogonal $W^\perp$ de W.

Queremos encontrar todos os vetores $w=(x, y, z, s, t)$ tais que

$$\begin{aligned}\langle w,u\rangle &= x+2y+3z-s+2t=0\\ \langle w,v\rangle &= 2x+4y+7z+2s-t=0\end{aligned}$$

Eliminando x da segunda equação, obtemos o sistema equivalente

$$\begin{aligned}x+2y+3z-s+2t&=0\\ z+4s-5t&=0\end{aligned}$$

As variáveis livres são y, s e t. Logo,

(1) Tomando $y = -1$, $s = 0$, $t = 0$, obtemos a solução $w_1 = (2, -1, 0, 0, 0)$.

(2) Tomando $y = 0$, $s = 1$, $t = 0$, obtemos a solução $w_2 = (13, 0, -4, 1, 0)$.

(3) Tomando $y = 0$, $s = 0$, $t = 1$, obtemos a solução $w_3 = (-17, 0, 5, 0, 1)$.

O conjunto $\{w_1, w_2, w_3\}$ é uma base de $W^{\perp}$.

7.12 Seja $w = (1, 2, 3, 1)$ um vetor de $\mathbf{R}^4$. Encontre uma base ortogonal de $w^{\perp}$.

Tomemos uma solução não nula de $x + 2y + 3z + t = 0$, digamos, $v_1 = (0, 0, 1, -3)$. Agora, tomemos uma solução não nula do sistema

$$x + 2y + 3z + t = 0, \qquad z - 3t = 0$$

digamos, $v_2 = (0, -5, 3, 1)$. Por último, calculamos uma solução não nula do sistema

$$x + 2y + 3z + t = 0, \qquad -5y + 3z + t = 0, \qquad z - 3t = 0$$

digamos, $v_3 = (-14, 2, 3, 1)$. Assim, v_1, v_2, v_3 é uma base ortogonal de $w^{\perp}$.

7.13 Seja S o conjunto constituído dos vetores de $\mathbf{R}^4$ dados.

$$u_1 = (1, 1, 0, -1), \;\; u_2 = (1, 2, 1, 3), \;\; u_3 = (1, 1, -9, 2), \;\; u_4 = (16, -13, 1, 3)$$

(a) Mostre que S é ortogonal e uma base de $\mathbf{R}^4$.

(b) Encontre as coordenadas de um vetor $v = (a, b, c, d)$ arbitrário de $\mathbf{R}^4$ em relação à base S.

(a) Calculamos

$$u_1 \cdot u_2 = 1 + 2 + 0 - 3 = 0, \qquad u_1 \cdot u_3 = 1 + 1 + 0 - 2 = 0, \qquad u_1 \cdot u_4 = 16 - 13 + 0 - 3 = 0$$
$$u_2 \cdot u_3 = 1 + 2 - 9 + 6 = 0, \qquad u_2 \cdot u_4 = 16 - 26 + 1 + 9 = 0, \qquad u_3 \cdot u_4 = 16 - 13 - 9 + 6 = 0$$

Assim, S é ortogonal e S é linearmente independente. Por isso, S é uma base de $\mathbf{R}^4$, já que quaisquer quatro vetores linearmente independentes de $\mathbf{R}^4$ formam uma base de $\mathbf{R}^4$.

(b) Como S é ortogonal, basta encontrar os coeficientes de Fourier de v em relação aos vetores da base, como no Teorema 7.7 Assim,

$$k_1 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{a + b - d}{3}, \qquad k_3 = \frac{\langle v, u_3\rangle}{\langle u_3, u_3\rangle} = \frac{a + b - 9c + 2d}{87}$$
$$k_2 = \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{a + 2b + c + 3d}{15}, \qquad k_4 = \frac{\langle v, u_4\rangle}{\langle u_4, u_4\rangle} = \frac{16a - 13b + c + 3d}{435}$$

são as coordenadas de v em relação à base S.

7.14 Sejam S, S_1, S_2 subespaços de V. Demonstre as afirmações seguintes.

(a) $S \subseteq S^{\perp\perp}$.

(b) Se $S_1 \subseteq S_2$, então $S_2^{\perp} \subseteq S_1^{\perp}$.

(c) $S^{\perp} = \text{ger }(S)^{\perp}$.

(a) Seja $w \in S$. Então $\langle w, v\rangle = 0$, para cada $v \in S^{\perp}$, portanto, $w \in S^{\perp\perp}$. Assim, $S \subseteq S^{\perp\perp}$.

(b) Seja $w \in S_2^{\perp}$. Então $\langle w, v\rangle = 0$, para cada $v \in S_2$. Como $S_1 \subseteq S_2$, $\langle w, v\rangle = 0$, temos $v = S_1$, para cada $v \in S_1$. Assim, $w \in S_1^{\perp}$ e, portanto, $S_2^{\perp} \subseteq S_1^{\perp}$.

(c) Como $S \subseteq \text{ger}(S)$, a parte (b) fornece $\text{ger}(S)^{\perp} \subseteq S^{\perp}$. Suponha que $u \in S^{\perp}$ e $v \in \text{ger}(S)$. Então existem $w_1, w_2, ..., w_k$ em S tais que $v = a_1w_1 + a_2w_2 + \cdots + a_kw_k$. Usando $u \in S^{\perp}$, temos

$$\begin{aligned}\langle u, v\rangle &= \langle u, \;\; a_1w_1 + a_2w_2 + \cdots + a_kw_k\rangle = a_1\langle u, w_1\rangle + a_2\langle u, w_2\rangle + \cdots + a_k\langle u, w_k\rangle \\ &= a_1(0) + a_2(0) + \cdots + a_k(0) = 0\end{aligned}$$

Assim, $u \in \text{ger}(S)^{\perp}$. Segue que $S^{\perp} \subseteq \text{ger}(S)^{\perp}$. Ambas inclusões fornecem $S^{\perp} = \text{ger}(S)^{\perp}$.

7.15 Demonstre o Teorema 7.5. Se S é um conjunto ortogonal de vetores não nulos, então S é linearmente independente.

Suponha que $S = \{u_1, u_2, ..., u_r\}$ e que

$$a_1u_1 + a_2u_2 + \cdots + a_ru_r = 0 \qquad (1)$$

Tomando o produto interno de (1) com u_1,

$$\begin{aligned} 0 = \langle 0, u_1\rangle &= \langle a_1u_1 + a_2u_2 + \cdots + a_ru_r,\ \ u_1\rangle \\ &= a_1\langle u_1, u_1\rangle + a_2\langle u_2, u_1\rangle + \cdots + a_r\langle u_r, u_1\rangle \\ &= a_1\langle u_1, u_1\rangle + a_2\cdot 0 + \cdots + a_r\cdot 0 = a_1\langle u_1, u_1\rangle \end{aligned}$$

Como $u_1 \neq 0$, resulta $\langle u_1, u_1\rangle \neq 0$ e, assim, $a_1 = 0$. Analogamente, para $i = 2, ..., n$, tomando o produto interno de (1) com u_i,

$$\begin{aligned} 0 = \langle 0, u_i\rangle &= \langle a_1u_1 + \cdots + a_ru_r,\ \ u_i\rangle \\ &= a_1\langle u_1, u_i\rangle + \cdots + a_i\langle u_i, u_i\rangle + \cdots + a_r\langle u_r, u_i\rangle = a_i\langle u_i, u_i\rangle \end{aligned}$$

Mas $\langle u_i, u_i\rangle \neq 0$ e, portanto, cada $a_i = 0$. Assim, S é linearmente independente.

7.16 Demonstre o Teorema 7.6 (Pitágoras). Seja $\{u_1, u_2, ..., u_r\}$ um conjunto ortogonal de vetores. Então

$$\|u_1 + u_2 + \cdots + u_r\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_r\|^2$$

Expandindo o produto interno, obtemos

$$\begin{aligned} \|u_1 + u_2 + \cdots + u_r\|^2 &= \langle u_1 + u_2 + \cdots + u_r,\ \ u_1 + u_2 + \cdots + u_r\rangle \\ &= \langle u_1, u_1\rangle + \langle u_2, u_2\rangle + \cdots + \langle u_r, u_r\rangle + \sum_{i\neq j}\langle u_i, u_j\rangle \end{aligned}$$

O teorema decorre observando que $\langle u_i, u_i\rangle = \|u_i\|^2$ e $\langle u_i, u_j\rangle = 0$, para $i \neq j$.

7.17 Demonstre o Teorema 7.7. Seja $\{u_1, u_2, ..., u_n\}$ uma base ortogonal de V. Então, dado qualquer $v \in V$,

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle}u_n$$

Suponha que $v = k_1u_1 + k_2u_2 + \cdots + k_nu_n$. Tomando o produto interno de ambos lados com u_1, obtemos

$$\begin{aligned} \langle v, u_1\rangle &= \langle k_1u_2 + k_2u_2 + \cdots + k_nu_n,\ \ u_1\rangle \\ &= k_1\langle u_1, u_1\rangle + k_2\langle u_2, u_1\rangle + \cdots + k_n\langle u_n, u_1\rangle \\ &= k_1\langle u_1, u_1\rangle + k_2\cdot 0 + \cdots + k_n\cdot 0 = k_1\langle u_1, u_1\rangle \end{aligned}$$

Assim, $k_1 = \dfrac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}$. Analogamente, para $i = 2, ..., n$,

$$\begin{aligned} \langle v, u_i\rangle &= \langle k_1u_i + k_2u_2 + \cdots + k_nu_n,\ \ u_i\rangle \\ &= k_1\langle u_1, u_i\rangle + k_2\langle u_2, u_i\rangle + \cdots + k_n\langle u_n, u_i\rangle \\ &= k_1\cdot 0 + \cdots + k_i\langle u_i, u_i\rangle + \cdots + k_n\cdot 0 = k_i\langle u_i, u_i\rangle \end{aligned}$$

Assim, $k_i = \dfrac{\langle v, u_i\rangle}{\langle u_1, u_i\rangle}$. Substituindo essa expressão no lugar de k_i na equação $v = k_1u_1 + \cdots + k_nu_n$, obtemos o resultado.

7.18 Seja $\{e_1, e_2, ..., e_n\}$ uma base ortonormal de V. Demonstre.

(a) Dado qualquer $u \in V$, temos $u = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \cdots + \langle u, e_n\rangle e_n$.

(b) $\langle a_1e_1 + \cdots + a_ne_n,\ \ b_1e_1 + \cdots + b_ne_n\rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

(c) Dados quaisquer $u, v \in V$, temos $\langle u, v\rangle = \langle u, e_1\rangle\langle v, e_1\rangle + \cdots + \langle u, e_n\rangle\langle v, e_n\rangle$.

(a) Suponha que $u = k_1e_1 + k_2e_2 + \cdots + k_ne_n$. Tomando o produto interno de u com e_1,

$$\begin{aligned} \langle u, e_1\rangle &= \langle k_1e_1 + k_2e_2 + \cdots + k_ne_n,\ \ e_1\rangle \\ &= k_1\langle e_1, e_1\rangle + k_2\langle e_2, e_1\rangle + \cdots + k_n\langle e_n, e_1\rangle \\ &= k_1(1) + k_2(0) + \cdots + k_n(0) = k_1 \end{aligned}$$

Analogamente, para $i = 2, 3, ..., n$,

$$\begin{aligned}\langle u, e_i\rangle &= \langle k_1e_1 + \cdots + k_ie_i + \cdots + k_ne_n, \quad e_i\rangle \\ &= k_1\langle e_1, e_i\rangle + \cdots + k_i\langle e_i, e_i\rangle + \cdots + k_n\langle e_n, e_i\rangle \\ &= k_1(0) + \cdots + k_i(1) + \cdots + k_n(0) = k_i\end{aligned}$$

Substituindo $\langle u, e_i\rangle$ no lugar de k_i na equação $u = k_1e_1 + \cdots + k_ne_n$, obtemos o resultado.

(b) Temos

$$\left\langle \sum_{i=1}^{n} a_ie_i, \quad \sum_{j=1}^{n} b_je_j \right\rangle = \sum_{i,j=1}^{n} a_ib_j\langle e_i, e_j\rangle = \sum_{i=1}^{n} a_ib_i\langle e_i, e_i\rangle + \sum_{i\neq j} a_ib_j\langle e_i, e_j\rangle$$

Mas $\langle e_i, e_j\rangle = 0$, para $i \neq j$, e $\langle e_i, e_j\rangle = 1$, para $i = j$. Logo, como queríamos,

$$\left\langle \sum_{i=1}^{n} a_ie_i, \quad \sum_{j=1}^{n} b_je_j \right\rangle = \sum_{i=1}^{n} a_ib_i = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

(c) Pela parte (a), temos

$$u = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n \qquad \text{e} \qquad v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$$

Assim, pela parte (b),

$$\langle u, v\rangle = \langle u, e_1\rangle\langle v, e_1\rangle + \langle u, e_2\rangle\langle v, e_2\rangle + \cdots + \langle u, e_n\rangle\langle v, e_n\rangle$$

Projeções, algoritmo de Gram-Schmidt, aplicações

7.19 Suponha que $w \neq 0$. Seja v um vetor qualquer de V. Mostre que

$$c = \frac{\langle v, w\rangle}{\langle w, w\rangle} = \frac{\langle v, w\rangle}{\|w\|^2}$$

é o único escalar tal que $v' = v - cw$ é ortogonal a w.

Para ter v' ortogonal a w, devemos ter

$$\langle v - cw, \quad w\rangle = 0 \qquad \text{ou} \qquad \langle v, w\rangle - c\langle w, w\rangle = 0 \qquad \text{ou} \qquad \langle v, w\rangle = c\langle w, w\rangle$$

Assim, $c \dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. Reciprocamente, suponha que $c = \dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. Então

$$\langle v - cw, \quad w\rangle = \langle v, w\rangle - c\langle w, w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w \ \ w\rangle}\langle w, w\rangle = 0$$

7.20 Encontre o coeficiente de Fourier c e a projeção de $v = (1, –2, 3, –4)$ sobre $w = (1, 2, 1, 2)$ em $\mathbf{R}^4$.

Calculamos $\langle v, w\rangle = 1 - 4 + 3 - 8 = -8$ e $\|w\|^2 = 1 + 4 + 1 + 4 = 10$. Então

$$c = -\tfrac{8}{10} = -\tfrac{4}{5} \qquad \text{e} \qquad \text{proj}(v, w) = cw = (-\tfrac{4}{5}, -\tfrac{8}{5}, -\tfrac{4}{5}, -\tfrac{8}{5})$$

7.21 Considere o subespaço U de $\mathbf{R}^4$ gerado pelos vetores

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, 1, 2, 4), \qquad v_3 = (1, 2, -4, -3)$$

Encontre (a) uma base ortogonal de U; (b) uma base ortonormal de U.

(a) Usamos o algoritmo de Gram-Schmidt. Começamos denotando $w_1 = u = (1, 1, 1, 1)$. Em seguida, calculamos

$$v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}w_1 = (1, 1, 2, 4) - \frac{8}{4}(1, 1, 1, 1) = (-1, -1, 0, 2)$$

Denotamos $w_2 = (-1, -1, 0, 2)$ e calculamos

$$\begin{aligned}v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle}w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle}w_2 &= (1, 2, -4, -3) - \frac{(-4)}{4}(1, 1, 1, 1) - \frac{(-9)}{6}(-1, -1, 0, 2) \\ &= (\tfrac{1}{2}, \tfrac{3}{2}, -3, 1)\end{aligned}$$

Limpamos as frações para obter $w_3 = (1, 3, -6, 2)$. Então w_1, w_2, w_3 formam uma base ortogonal de U.

(b) Normalizamos a base ortogonal formada por w_1, w_2, w_3. Como $\|w_1\|^2 = 4$, $\|w_2\|^2 = 6$ e $\|w_3\|^2 = 50$, o conjunto de vetores seguintes forma uma base ortonormal de U.

$$u_1 = \frac{1}{2}(1, 1, 1, 1), \qquad u_2 = \frac{1}{\sqrt{6}}(-1, -1, 0, 2), \qquad u_3 = \frac{1}{5\sqrt{2}}(1, 3, -6, 2)$$

7.22 Considere o espaço vetorial $\mathbf{P}(t)$ com o produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$. Aplique o algoritmo de Gram-Schmidt ao conjunto $\{1, t, t^2\}$ para obter um conjunto ortogonal $\{f_0, f_1, f_2\}$ com coeficientes inteiros.

Começamos denotando $f_0 = 1$. Em seguida, calculamos

$$t - \frac{\langle t, 1\rangle}{\langle 1, 1\rangle}\cdot 1 = t - \frac{\frac{1}{2}}{1}\cdot 1 = t - \frac{1}{2}$$

Limpamos as frações para obter $f_1 = 2t - 1$. Então calculamos

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^2,\ 2t-1\rangle}{\langle 2t-1,\ 2t-1\rangle}(2t-1) = t^2 - \frac{\frac{1}{3}}{1}(1) - \frac{\frac{1}{6}}{\frac{1}{3}}(2t-1) = t^2 - t + \frac{1}{6}$$

Limpamos as frações para obter $f_2 = 6t^2 - 6t + 1$. Assim, $\{1,\ 2t-1,\ 6t^2 - 6t + 1\}$ é o conjunto ortogonal procurado.

7.23 Seja $v = (1, 3, 5, 7)$. Encontre a projeção de v sobre W ou, em outras palavras, encontre $w \in W$ que minimiza $\|v - w\|$, sendo W o subespaço de $\mathbf{R}^4$ gerado por

(a) $u_1 = (1, 1, 1, 1)$ e $u_2 = (1, -3, 4, -2)$,

(b) $v_1 = (1, 1, 1, 1)$ e $v_2 = (1, 2, 3, 2)$.

(a) Como u_1 e u_2 são ortogonais, basta calcular os coeficientes de Fourier

$$c_1 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{1+3+5+7}{1+1+1+1} = \frac{16}{4} = 4$$

$$c_2 = \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{1-9+20-14}{1+9+16+4} = \frac{-2}{30} = -\frac{1}{15}$$

Então $w = \text{proj}(v, W) = c_1u_1 + c_2u_2 = 4(1, 1, 1, 1) - \frac{1}{15}(1, -3, 4, -2) = (\frac{59}{15}, \frac{63}{5}, \frac{56}{15}, \frac{62}{15})$.

(b) Como u_1 e u_2 não são ortogonais, primeiro aplicamos o algoritmo de Gram-Schmidt para encontrar uma base ortogonal de W. Denotando $w_1 = v_1 = (1, 1, 1, 1)$, encontramos

$$v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}w_1 = (1, 2, 3, 2) - \frac{8}{4}(1, 1, 1, 1) = (-1, 0, 1, 0)$$

Denotando $w_2 = (-1, 0, 1, 0)$, calculamos

$$c_1 = \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle} = \frac{1+3+5+7}{1+1+1+1} = \frac{16}{4} = 4$$

$$c_2 = \frac{\langle v, w_2\rangle}{\langle w_2, w_2\rangle} - \frac{-1+0+5+0}{1+0+1+0} = \frac{-6}{2} = -3$$

Então $w = \text{proj}(v, W) = c_1w_1 + c_2w_2 = 4(1, 1, 1, 1) - 3(-1, 0, 1, 0) = (7, 4, 1, 4)$.

7.24 Sejam w_1 e w_2 vetores não nulos ortogonais e v um vetor qualquer de V. Encontre c_1 e c_2 tais que v' seja ortogonal a w_1 e w_2, com $v' = v - c_1w_1 - c_2w_2$.

Se v' é ortogonal a w_1, então

$$\begin{aligned} 0 &= \langle v - c_1w_1 - c_2w_2,\ w_1\rangle = \langle v, w_1\rangle - c_1\langle w_1, w_1\rangle - c_2\langle w_2, w_1\rangle \\ &= \langle v, w_1\rangle - c_1\langle w_1, w_1\rangle - c_2 0 = \langle v, w_1\rangle - c_1\langle w_1, w_1\rangle \end{aligned}$$

Assim, $c_1 = \langle v, w_1\rangle/\langle w_1, w_1\rangle$. (Ou seja, c_1 é o componente de v ao longo de w_1.) Analogamente, se v' é ortogonal a w_2, então

$$0 = \langle v - c_1w_1 - c_2w_2,\ w_2\rangle = \langle v, w_2\rangle - c_2\langle w_2, w_2\rangle$$

Assim, $c_2 = \langle v, w_2\rangle/\langle w_2, w_2\rangle$. (Ou seja, c_2 é o componente de v ao longo de w_2.)

7.25 Demonstre o Teorema 7.8. Sejam $w_1, w_2, ..., w_r$ vetores ortogonais de V. Dado um vetor v qualquer de V, denotamos

$$v' = v - (c_1w_1 + c_2w_2 + \cdots + c_rw_r), \qquad \text{com} \qquad c_i = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}$$

Então v' é ortogonal a $w_1, w_2, ..., w_r$.

Dado $i = 1, 2, ..., r$ e usando que $\langle w_i, w_j\rangle = 0$, com $i \neq j$, temos

$$\begin{aligned}\langle v - c_1w_1 - c_2x_2 - \cdots - c_rw_r, \ \ w_i\rangle &= \langle v, w_i\rangle - c_1\langle w_1, w_i\rangle - \cdots - c_i\langle w_i, w_i\rangle - \cdots - c_r\langle w_r, w_i\rangle \\ &= \langle v, w_i\rangle - c_1 \cdot 0 - \cdots - c_i\langle w_i, w_i\rangle - \cdots - c_r \cdot 0 \\ &= \langle v, w_i\rangle - c_i\langle w_i, w_i\rangle = \langle v, w_i\rangle - \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle = 0\end{aligned}$$

Isso demonstra o teorema.

7.26 Demonstre o Teorema 7.9. Seja $\{v_1, v_2, ..., v_n\}$ uma base qualquer de um espaço com produto interno V. Então existe uma base ortonormal $\{u_1, u_2, ..., u_n\}$ de V tal que a matriz de mudança de base de $\{v_i\}$ para $\{u_i\}$ é triangular, ou seja, para $k = 1, 2, ..., n$,

$$u_k = a_{k1}v_1 + a_{k2}v_2 + \cdots + a_{kk}v_k$$

A demonstração utiliza o algoritmo de Gram-Schmidt e as observações 1 e 3 da Seção 7.7. Mais especificamente, aplicamos o algoritmo a $\{v_i\}$ para obter uma base ortogonal $\{w_1, ..., w_n\}$, que normalizamos para obter uma base ortonormal $\{u_i\}$ de V. O algoritmo especificamente garante que cada w_k é uma combinação linear de $v_1, ..., v_n$ e, portanto, cada u_k é uma combinação linear de $v_1, ..., v_n$.

7.27 Demonstre o Teorema 7.10. Seja $S = \{w_1, w_2, ..., w_r\}$ uma base ortogonal de um subespaço W de um espaço vetorial V. Então S pode ser estendida a uma base ortogonal de V, ou seja, podemos encontrar vetores $w_{r+1}, \ldots, w_r$ tais que $\{w_1, w_2, ..., w_n\}$ seja uma base ortogonal de V.

Estendemos S a uma base $S' = \{w_1, \ldots, w_r, v_{r+1}, \ldots, v_n\}$ de V. Aplicando o algoritmo de Gram-Schmidt a S', primeiro obtemos $w_1, ..., w_r$, porque S é ortogonal, e depois obtemos vetores $w_{r+1}, \ldots, w_n$ tais que $\{w_1, w_2, ..., w_n\}$ é uma base ortogonal de V. Assim, provamos o teorema.

7.28 Demonstre o Teorema 7.4. Seja W um subespaço de V. Então $V = W \oplus W^{\perp}$.

Pelo Teorema 7.9, existe uma base ortogonal $\{u_1, ..., u_r\}$ de W e, pelo Teorema 7.10, podemos estendê-la a uma base ortogonal $\{u_1, u_2, ..., u_n\}$ de V. Logo, $u_{r+1}, \ldots, u_n \in W^{\perp}$. Dado $v \in V$, temos

$$v = a_1u_1 + \cdots + a_nu_n, \text{ com } a_1u_1 + \cdots + a_ru_r \in W \text{ e } a_{r+1}u_{r+1} + \cdots + a_nu_n \in W^{\perp}$$

Assim, obtemos $V = W + W^{\perp}$.

Por outro lado, se $w \in W \cap W^{\perp}$, então $\langle w, w\rangle = 0$, fornecendo $w = 0$. Logo, $W \cap W^{\perp} = \{0\}$.

As duas condições $V = W + W^{\perp}$ e $W \cap W^{\perp} = \{0\}$ dão o resultado procurado $V = W \oplus W^{\perp}$.

OBSERVAÇÃO Nossa demonstração foi feita no caso em que V tem dimensão finita. Observamos que o teorema também é válido em espaços de dimensão infinita.

7.29 Seja W um subespaço de um espaço V de dimensão finita. Prove que $W = W^{\perp\perp}$.

Pelo Teorema 7.4, $V = W \oplus W^{\perp}$ e, também $V = W^{\perp} \oplus W^{\perp\perp}$. Logo,

$$\dim W = \dim V - \dim W^{\perp} \qquad \text{e} \qquad \dim W^{\perp\perp} = \dim V - \dim W^{\perp}$$

Isso fornece $\dim W = \dim W^{\perp\perp}$. Como $W \subseteq W^{\perp\perp}$ (ver Problema 7.14), resulta $W = W^{\perp\perp}$.

7.30 Demonstre a afirmação dada. Seja $w_1, w_2, ..., w_r$ um conjunto ortogonal de vetores não nulos de V. Seja v um vetor qualquer de V e c_i o componente de v ao longo de w_i. Dados escalares $a_1, ..., a_r$ quaisquer, temos

$$\left\| v - \sum_{k=1}^{r} c_kw_k \right\| \leq \left\| v - \sum_{k=1}^{r} a_kw_k \right\|$$

Ou seja, $\sum c_iw_i$ é a melhor aproximação de v por combinações lineares de $w_1, ..., w_r$.

Pelo Teorema 7.8, $v - \sum c_kw_k$ é ortogonal a cada w_i e, portanto, ortogonal a qualquer combinação linear dos vetores $w_1, w_2, ..., w_r$. Portanto, usando o teorema de Pitágoras e somando de $k = 1$ a r, obtemos

$$\|v - \sum a_k w_k\|^2 = \|v - \sum c_k w_k + \sum (c_k - a_k) w_k\|^2 = \|v - \sum c_k w_k\|^2 + \|\sum (c_k - a_k) w_k\|^2$$
$$\geq \|v - \sum c_k w_k\|^2$$

A raiz quadrada de ambos lados nos dá o teorema.

7.31 Seja $\{e_1, e_2, ..., e_r\}$ um conjunto ortonormal de vetores de V. Seja v um vetor qualquer de V e c_i o coeficiente de Fourier de v em relação a e_i. Demonstre a desigualdade de Bessel

$$\sum_{k=1}^{r} c_k^2 \leq \|v\|^2$$

Observe que $c_i = \langle v, e_i \rangle$, porque $\|e_i\| = 1$. Então, usando $\langle e_i, e_j \rangle = 0$, com $i \neq j$, e somando de $k = 1$ a r, obtemos

$$0 \leq \langle v - \sum c_k e_k, \quad v - \sum c_k, e_k \rangle = \langle v, v \rangle - 2\langle v, \quad \sum c_k e_k \rangle + \sum c_k^2 = \langle v, v \rangle - \sum 2c_k \langle v, e_k \rangle + \sum c_k^2$$
$$= \langle v, v \rangle - \sum 2c_k^2 + \sum c_k^2 = \langle v, v \rangle - \sum c_k^2$$

que dá a desigualdade de Besel.

Matrizes ortogonais

7.32 Encontre uma matriz ortogonal P cuja primeira linha seja $u_1 = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

Inicialmente, procuramos um vetor não nulo $w_2 = (x, y, z)$ que seja ortogonal a u_1, ou seja, para o qual

$$0 = \langle u_1, w_2 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{ou} \qquad x + 2y + 2z = 0$$

Uma tal solução é $w_2 = (0, 1, -1)$. Normalizamos w_2 para obter a segunda linha de P,

$$u_2 = (0, 1/\sqrt{2}, -1/\sqrt{2})$$

Em seguida, procuramos um vetor não nulo $w_3 = (x, y, z)$ que seja ortogonal a u_1 e, também, a u_2, ou seja, para o qual

$$0 = \langle u_1, w_3 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{ou} \qquad x + 2y + 2z = 0$$
$$0 = \langle u_2, w_3 \rangle = \frac{y}{\sqrt{2}} - \frac{y}{\sqrt{2}} = 0 \qquad \text{ou} \qquad y - z = 0$$

Tomando $z = -1$, encontramos a solução $w_3 = (4, -1, -1)$. Normalizamos w_3 para obter a terceira linha de P, a saber,

$$u_3 = (4/\sqrt{18}, -1/\sqrt{18}, -1/\sqrt{18}).$$

Assim, $P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 4/3\sqrt{2} & -1/3\sqrt{2} & -1/3\sqrt{2} \end{bmatrix}$

Enfatizamos que a matriz P encontrada não é única.

7.33 Seja $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 4 \\ 7 & -5 & 2 \end{bmatrix}$. Decida se é válido, ou não. (a) As linhas de A são ortogonais; (b) A é uma matriz ortogonal; (c) as colunas de A são ortogonais.

(a) É válido, porque $(1, 1, -1) \cdot (1, 3, 4) = 1 + 3 - 4 = 0$, $(1, 1 - 1) \cdot (7, -5, 2) = 7 - 5 - 2 = 0$ e $(1, 3, 4) \cdot (7, -5, 2) = 7 - 15 + 8 = 0$.

(b) Não é, porque as linhas de A não são vetores unitários; por exemplo, $\|(1, 1, -1)\|^2 = 1 + 1 + 1 = 3$.

(c) Não é; por exemplo, $(1, 1, 7) \cdot (1, 3, -5) = 1 + 3 - 35 = -31 \neq 0$.

7.34 Seja B a matriz obtida normalizando cada linha da matriz A do Problema 7.33.

(a) Encontre B.

(b) Será B ortogonal?

(c) Serão ortogonais as colunas de B?

(a) Temos

$$\|(1,1,-1)\|^2 = 1+1+1 = 3, \qquad \|(1,3,4)\|^2 = 1+9+16 = 26$$
$$\|(7,-5,2)\|^2 = 49+25+4 = 78$$

Assim,
$$B = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1/\sqrt{26} & 3/\sqrt{26} & 4/\sqrt{26} \\ 7/\sqrt{78} & -5/\sqrt{78} & 2/\sqrt{78} \end{bmatrix}$$

(b) É, porque as linhas de B continuam ortogonais e agora são vetores unitários.

(c) São, porque as linhas de B formam um conjunto ortonormal de vetores. Então, pelo Teorema 7.11, as colunas de B formam, automaticamente, um conjunto ortonormal.

7.35 Demonstre cada uma das afirmações seguintes.

(a) P é ortogonal se, e só se, P^T é ortogonal.

(b) Se P é ortogonal, então P^{-1} é ortogonal.

(c) Se P e Q são ortogonais, então PQ é ortogonal.

(a) Temos $(P^T)^T = P$. Assim, P é ortogonal se, e só se, $PP^T = I$ se, e só se, $P^{TT}P^T = I$ se, e só se, P^T é ortogonal.

(b) Temos $P^T = P^{-1}$, porque P é ortogonal. Assim, pela parte (a), P^{-1} é ortogonal.

(c) Temos $P^T = P^{-1}$ e $Q^T = Q^{-1}$. Assim, $(PQ)(PQ)^T = PQQ^TP^T = PQQ^{-1}P^{-1} = I$. Logo, $(PQ)^T = (PQ)^{-1}$ e, portanto PQ é ortogonal.

7.36 Seja P uma matriz ortogonal. Mostre que

(a) $\langle Pu, Pv\rangle = \langle u, v\rangle$, para quaisquer $u, v \in V$;

(b) $\|Pu\| = \|u\|$, para cada $u \in V$.

Usamos $P^TP = I$ e $\langle u, v\rangle = u^Tv$.

(a) $\langle Pu, Pv\rangle = (Pu)^T(Pv) = u^TP^TPv = u^Tv = \langle u, v\rangle$.

(b) Temos

$$\|Pu\|^2 = \langle Pu, Pu\rangle = u^TP^TPu = u^Tu = \langle u, u\rangle = \|u\|^2$$

Tomando a raiz quadrada de ambos lados, obtemos (b).

7.37 Demonstre o Teorema 7.12. Sejam $E = \{e_i\}$ e $E' = \{e'_i\}$ bases ortonormais de V e P a matriz de mudança de base da base E para a base E'. Então P é ortogonal.

Suponha que

$$e'_i = b_{i1}e_1 + b_{i2}e_2 + \cdots + b_{in}e_n, \qquad i = 1, \ldots, n \tag{1}$$

Usando o Problema 7.18(b) e a hipótese da ortonormalidade de E', obtemos

$$\delta_{ij} = \langle e'_i, e'_j\rangle = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \tag{2}$$

Seja $B = [b_{ij}]$ a matriz de coeficientes de (1). (Logo, $P = B^T$.) Suponha que $BB^T = [c_{ij}]$. Então

$$c_{ij} = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \tag{3}$$

Por (2) e (3), temos $c_{ij} = \delta_{ij}$. Assim, $BB^T = I$. Decorre que B é ortogonal e, portanto, $P = B^T$ é ortogonal.

7.38 Demonstre o Teorema 7.13. Sejam $\{e_1, ..., e_n\}$ uma base ortonormal de um espaço com produto interno V e $P = [a_{ij}]$ uma matriz ortogonal. Então os n vetores seguintes formam uma base ortonormal de V.

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, 2, \ldots, n$$

Como $\{e_i\}$ é ortonormal, o Problema 7.18(b) garante que

$$\langle e'_i, e'_j\rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \langle C_i, C_j\rangle$$

onde C_i denota a i-ésima coluna da matriz ortogonal $P = [a_{ij}]$. Como P é ortogonal, suas colunas formam um conjunto ortonormal. Isso implica $\langle e_i', e_j' \rangle = \langle C_i, C_j \rangle = \delta_{ij}$. Assim, $\{e_i'\}$ é uma base ortonormal.

Produtos internos e matrizes positivas

7.39 Decida quais das matrizes simétricas dadas são positivas.

(a) $A = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 8 & -3 \\ -3 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$, (d) $D = \begin{bmatrix} 3 & 5 \\ 5 & 9 \end{bmatrix}$

Usamos o Teorema 7.14, que afirma que uma matriz 2×2 real e simétrica é positiva se, e só se, suas entradas diagonais são positivas e seu determinante é positivo.

(a) Não é, porque $|A| = 15 - 16 = -1$ é negativo.

(b) É positiva.

(c) Não é, porque a entrada diagonal –3 é negativa.

(d) É positiva.

7.40 Encontre os valores de k com os quais as matrizes dadas são positivas.

(a) $A = \begin{bmatrix} 2 & -4 \\ -4 & k \end{bmatrix}$, (b) $B = \begin{bmatrix} 4 & k \\ k & 9 \end{bmatrix}$, (c) $C = \begin{bmatrix} k & 5 \\ 5 & -2 \end{bmatrix}$

(a) Em primeiro lugar, k deve ser positivo. Também $|A| = 2k - 16$ deve ser positivo, portanto, $2k - 16 > 0$. Assim, $k > 8$.

(b) Precisamos que $|B| = 36 - k^2$ seja positivo, ou seja, $36 - k^2 > 0$. Assim, $k^2 < 36$, ou $-6 < k < 6$.

(c) C nunca será positiva, pois tem a entrada negativa –2 na diagonal.

7.41 Encontre a matriz A que representa o produto interno canônico de $\mathbf{R}^2$ em relação a cada uma das bases dadas de $\mathbf{R}^2$. (a) $\{v_1 = (1,3),\ v_2 = (2,5)\}$; (b) $\{w_1 = (1,2),\ w_2 = (4,-2)\}$.

(a) Calculamos $\langle v_1, v_1 \rangle = 1 + 9 = 10$, $\langle v_1, v_2 \rangle = 2 + 15 = 17$, $\langle v_2, v_2 \rangle = 4 + 25 = 29$. Assim,

$$A = \begin{bmatrix} 10 & 17 \\ 17 & 29 \end{bmatrix}$$

(b) Calculamos $\langle w_1, w_1 \rangle = 1 + 4 = 5$, $\langle w_1, w_2 \rangle = 4 - 4 = 0$, $\langle w_2, w_2 \rangle = 16 + 4 = 20$. Assim, $A = \begin{bmatrix} 5 & 0 \\ 0 & 20 \end{bmatrix}$

(Como os vetores da base são ortogonais, a matriz A é diagonal.)

7.42 Considere o espaço vetorial $V = \mathbf{P}_2(t)$ com produto interno $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\, dt$.

(a) Encontre $\langle f, g \rangle$, com $f(t) = t + 2$ e $g(t) = t^2 - 3t + 4$.

(b) Encontre a matriz A do produto interno em relação à base $\{1, t, t^2\}$ de V.

(c) Verifique o Teorema 7.16, mostrando que $\langle f, g \rangle = [f]^T A[g]$ em relação à base $\{1, t, t^2\}$.

(a) $\langle f, g \rangle = \int_{-1}^{1} (t+2)(t^2 - 3t + 4)\, dt = \int_{-1}^{1} (t^3 - t^2 - 2t + 8)\, dt = \left(\frac{t^4}{4} - \frac{t^3}{3} - t^2 + 8t \right) \Big|_{-1}^{1} = \frac{46}{3}$

(b) Aqui usamos o fato de que, se $r + s = n$, então

$$\langle t^r, t^r \rangle = \int_{-1}^{1} t^n\, dt = \frac{t^{n+1}}{n+1} \Big|_{-1}^{1} = \begin{cases} 2/(n+1) & \text{se } n \text{ é par,} \\ 0 & \text{se } n \text{ é ímpar.} \end{cases}$$

Então $\langle 1, 1 \rangle = 2$, $\langle 1, t \rangle = 0$, $\langle 1, t^2 \rangle = \frac{2}{3}$, $\langle t, t \rangle = \frac{2}{3}$, $\langle t, t^2 \rangle = 0$, $\langle t^2, t^2 \rangle = \frac{2}{5}$. Assim,

$$A = \begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix}$$

(c) Temos $[f]^T = (2, 1, 0)$ e $[g]^T = (4, -3, 1)$ em relação à base dada. Então

$$[f]^T A[g] = (2, 1, 0)\begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix}\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix} = (4, \tfrac{2}{3}, \tfrac{4}{3})\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix} = \tfrac{46}{3} = \langle f, g\rangle$$

7.43 Demonstre o Teorema 7.14. $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ é positiva se, e só se, a e d são positivos e $|A| = ad - b^2$ é positivo.

Seja $u = [x, y]^T$. Então

$$f(u) = u^T A u = [x, y]\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + dy^2$$

Suponha que $f(u) > 0$, para cada $u \neq 0$. Então, $f(1, 0) = a > 0$ e $f(0, 1) = d > 0$. Também, $f(b, -a) = a(ad - b^2) > 0$. Como $a > 0$, obtemos $ad - b^2 > 0$.

Reciprocamente, suponha que $a > 0$, $d > 0$, $ad - b^2 > 0$. Completando o quadrado, obtemos

$$f(u) = a\left(x^2 + \frac{2b}{a}xy + \frac{b^2}{a^2}y^2\right) + dy^2 - \frac{b^2}{a}y^2 = a\left(x + \frac{by}{a}\right)^2 + \frac{ad - b^2}{a}y^2$$

Dessa forma, $f(u) > 0$, para cada $u \neq 0$.

7.44 Demonstre o Teorema 7.15. Seja A uma matriz real positiva. Então a função $\langle u, v\rangle = u^T A v$ é um produto interno de $\mathbf{R}^n$.

Dados vetores u_1, u_2 e v,

$$\langle u_1 + u_2,\ v\rangle = (u_1 + u_2)^T A v = (u_1^T + u_2^T)Av = u_1^T A v + u_2^T A v = \langle u_1, v\rangle + \langle u_2, v\rangle$$

e, dado qualquer escalar r e vetores u, v,

$$\langle ru, v\rangle = (ru)^T A v = r u^T A v = r\langle u, v\rangle$$

Logo, vale [I_1].

Como $u^T A v$ é um escalar, $(u^T A v)^T = u^T A v$. Também, $A^T = A$, pois A é simétrica. Logo,

$$\langle u, v\rangle = u^T A v = (u^T A v)^T = v^T A^T u^{TT} = v^T A u = \langle v, u\rangle$$

Logo, vale [I_2].
Finalmente, como A é positiva, temos $X^T A X > 0$, para cada $X \in \mathbf{R}^n$ não nulo. Assim, dado qualquer vetor não nulo v, temos v, $\langle v, v\rangle = v^T A v > 0$. Também, $\langle 0, 0\rangle = 0^T A 0 = 0$. Assim, vale [$I_3$]. Em vista disso, a função $\langle u, v\rangle = Av$ é um produto interno.

7.45 Demonstre o Teorema 7.16. Seja A a representação matricial de um produto interno em relação a uma base S de V. Então, dados quaisquer vetores $u, v \in V$, temos

$$\langle u, v\rangle = [u]^T A[v]$$

Suponha que $S = \{w_1, w_2, ..., w_r\}$ e que $A = [k_{ij}]$. Então, $k_{ij} = \langle w_i, w_j\rangle$. Suponha que

$$u = a_1 w_1 + a_2 w_2 + \cdots + a_n w_n \qquad \text{e} \qquad v = b_1 w_1 + b_2 w_2 + \cdots + b_n w_n$$

Então

$$\langle u, v\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \langle w_i, w_j\rangle \tag{1}$$

Por outro lado,

$$[u]^T A[v] = (a_1, a_2, \dots, a_n)\begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ \dots & \dots & \dots & \dots \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= \left(\sum_{i=1}^{n} a_i k_{i1}, \ \sum_{i=1}^{n} a_i k_{i2}, \ \dots, \ \sum_{i=1}^{n} a_i k_{in}\right)\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{j=1}^{n}\sum_{i=1}^{n} a_i b_j k_{ij} \tag{2}$$

As equações (1) e (2) garantem a afirmação do teorema.

7.46 Demonstre o Teorema 7.17. Seja A a representação matricial de algum produto interno de V. Então A é uma matriz positiva.

Como $\langle w_i, w_j \rangle = \langle w_j, w_i \rangle$ vale para quaisquer vetores w_i e w_j da base, a matriz A é simétrica. Seja X um vetor não nulo qualquer de $\mathbf{R}^n$. Então $[u] = X$, para algum vetor $u \in V$ não nulo. O Teorema 7.16 afirma que $X^T AX = [u]^T A[u] = \langle u, u \rangle > 0$. Assim, A é positiva.

Espaços com produto interno complexos

7.47 Seja V um espaço com produto interno complexo. Verifique a relação

$$\langle u, \ av_1 + bv_2 \rangle = \bar{a}\langle u, v_1 \rangle + \bar{b}\langle u, v_2 \rangle$$

Usando $[I_2^*]$, $[I_1^*]$ e, depois, $[I_2^*]$, obtemos

$$\langle u, av_1 + bv_2 \rangle = \overline{\langle av_1 + bv_2, u \rangle} = \overline{a\langle v_1, u \rangle + b\langle v_2, u \rangle} = \bar{a}\overline{\langle v_1, u \rangle} + \bar{b}\overline{\langle v_2, u \rangle} = \bar{a}\langle u, v_1 \rangle + \bar{b}\langle u, v_2 \rangle$$

7.48 Suponha que $\langle u, v \rangle = 3 + 2i$ num espaço com produto interno complexo. Encontre

(a) $\langle (2-4i)u, v \rangle$; (b) $\langle u, \ (4+3i)v \rangle$; (c) $\langle (3-6i)u, \ (5-2i)v \rangle$.

(a) $\langle (2-4i)u, \ v \rangle = (2-4i)\langle u, v \rangle = (2-4i)(3+2i) = 14 - 8i$

(b) $\langle u, \ (4+3i)v \rangle = \overline{(4+3i)}\langle u, v \rangle = (4-3i)(3+2i) = 18 - i$

(c) $\langle (3-6i)u, \ (5-2i)v \rangle = (3-6i)\overline{(5-2i)}\langle u, v \rangle = (3-6i)(5+2i)(3+2i) = 129 - 18i$

7.49 Encontre o coeficiente de Fourier (componente) c e a projeção cw de $v = (3+4i, \ 2-3i)$ sobre $w = (5+i, \ 2i)$ em $\mathbf{C}^2$.

Lembrando que $c = \langle v, w \rangle / \langle w, w \rangle$, calculamos

$$\begin{aligned} \langle v, w \rangle &= (3+4i)(\overline{5+i}) + (2-3i)(\overline{2i}) = (3+4i)(5-i) + (2-3i)(-2i) \\ &= 19 + 17i - 6 - 4i = 13 + 13i \\ \langle w, w \rangle &= 25 + 1 + 4 = 30 \end{aligned}$$

Assim, $c = (13+13i)/30 = \frac{13}{30} + \frac{13}{30}i$. Decorre que $\text{proj}(v, w) = cw = (\frac{26}{15} + \frac{39}{15}i, \ -\frac{13}{15} + \frac{1}{15}i)$.

7.50 Demonstre o Teorema 7.18 (Cauchy-Schwarz). Seja V um espaço com produto interno complexo. Então $|\langle u, v \rangle| \le \|u\|\|v\|$.

Se $v = 0$, a desigualdade reduz a $0 \le 0$ e, portanto, é válida. Agora, suponha que $v \ne 0$. Usando $z\bar{z} = |z|^2$ (válido para números complexos) e $\langle v, u \rangle = \overline{\langle u, v \rangle}$, expandimos $\|u - \langle u, v \rangle tv\|^2 \ge 0$, em que t é um valor real.

$$\begin{aligned} 0 \le \|u - \langle u, v \rangle tv\|^2 &= \langle u - \langle u, v \rangle tv, \ \ u - \langle u, v \rangle tv \rangle \\ &= \langle u, u \rangle - \overline{\langle u, v \rangle} t \langle u, v \rangle - \langle u, v \rangle t \langle v, u \rangle + \langle u, v \rangle \overline{\langle u, v \rangle} t^2 \langle v, v \rangle \\ &= \|u\|^2 - 2t|\langle u, v \rangle|^2 + |\langle u, v \rangle|^2 t^2 \|v\|^2 \end{aligned}$$

Tomamos $t = 1/\|v\|^2$ para obter $0 \le \|u\|^2 - \dfrac{|\langle u, v\rangle|^2}{\|v\|^2}$, do que decorre $|\langle u, v\rangle|^2 \le \|v\|^2\|v\|^2$. Tomando a raiz quadrada de ambos lados, obtemos a desigualdade procurada.

7.51 Encontre uma base ortogonal de $u^\perp$ em C^3, com $u = (1,\ i,\ 1+i)$.

Aqui $u^\perp$ consiste em todos os vetores $s = (x, y, z)$ tais que

$$\langle w, u\rangle = x - iy + (1-i)z = 0$$

Encontramos uma solução, digamos, $w_1 = (0,\ 1-i,\ i)$. Então resolvemos o sistema

$$x - iy + (1-i)z = 0, \qquad (1+i)y - iz = 0$$

Aqui, z é uma variável livre. Tomamos $z = 1$ para obter $y = i/(1+i) = (1+i)/2$ e $x = (3i-3)2$. Multiplicando por 2, obtemos a solução $w_2 = (3i-3,\quad 1+i,\quad 2)$. Os vetores w_1 e w_2 formam uma base ortogonal de $u^\perp$.

7.52 Encontre uma base ortonormal do subespaço W de $\mathbf{C}^3$ gerado por

$$v_1 = (1, i, 0) \qquad \text{e} \qquad v_2 = (1,\ 2,\ 1-i).$$

Aplicamos o algoritmo de Gram-Schmidt. Tomamos $w_1 = v_1 = (1, i, 0)$ e calculamos

$$v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 = (1,\ 2,\ 1-i) - \frac{1-2i}{2}(1, i, 0) = (\tfrac{1}{2} + i,\ 1 - \tfrac{1}{2}i,\ 1-i)$$

Multiplicamos por 2 para limpar as frações, obtendo $w_2 = (1+2i,\ 2-i,\ 2-2i)$. Em seguida, calculamos $\|w_1\| = \sqrt{2}$ e $\|w_2\| = \sqrt{18}$. Normalizando $\{w_1, w_2\}$, obtemos a base ortonormal de W seguinte.

$$\left\{ u_1 = \left(\frac{1}{\sqrt{2}},\ \frac{i}{\sqrt{2}},\ 0\right),\ u_2 = \left(\frac{1+2i}{\sqrt{18}},\ \frac{2-i}{\sqrt{18}},\ \frac{2-2i}{\sqrt{18}}\right) \right\}$$

7.53 Encontre a matriz P que representa o produto interno canônico de $\mathbf{C}^3$ em relação à base $\{1,\ i,\ 1-i\}$.

Calculamos os seis produtos internos a seguir.

$$\langle 1, 1\rangle = 1, \qquad \langle 1, i\rangle = \bar{i} = -i, \qquad \langle 1, 1-i\rangle = \overline{1-i} = 1+i$$
$$\langle i, i\rangle = i\bar{i} = 1, \qquad \langle i, 1-i\rangle = i(\overline{1-i}) = -1+i, \qquad \langle 1-i, 1-i\rangle = 2$$

Em seguida, usando $(u, v) = \overline{\langle v, u\rangle}$, obtemos

$$P = \begin{bmatrix} 1 & -i & 1+i \\ i & 1 & -1+i \\ 1-i & -1-i & 2 \end{bmatrix}$$

(Conforme esperado, P é hermitiana, ou seja, $P^H = P$.)

Espaços vetoriais normados

7.54 Considere os vetores $u = (1, 3, -6, 4)$ e $v = (3, -5, 1, -2)$ de $\mathbf{R}^4$. Encontre
(a) $\|u\|_\infty$ e $\|v|_\infty$, (b) $\|u\|_1$ e $\|v\|_1$, (c) $\|u\|_2$ e $\|v\|_2$,
(d) $d_\infty(u, v), d_1(u, v), d_2(u, v)$.

(a) A norma infinito escolhe o máximo do valor absoluto dos componentes. Logo,

$$\|u\|_\infty = 6 \qquad \text{e} \qquad \|v\|_\infty = 5$$

(b) A norma 1 soma o valor absoluto dos componentes. Logo,

$$\|u\|_1 = 1+3+6+4 = 14 \qquad \text{e} \qquad \|v\|_1 = 3+5+1+2 = 11$$

(c) A norma 2 é igual à raiz quadrada da soma do quadrado dos componentes (ou seja, a norma induzida pelo produto interno canônico de $\mathbf{R}^3$). Logo,

$$\|u\|_2 = \sqrt{1+9+36+16} = \sqrt{62} \qquad \text{e} \qquad \|v\|_2 = \sqrt{9+25+1+4} = \sqrt{39}$$

(d) Primeiro calculamos $u - v = (-2, 8, -7, 6)$. Então

$$d_\infty(u, v) = \|u - v\|_\infty = 8$$
$$d_1(u, v) = \|u - v\|_1 = 2 + 8 + 7 + 6 = 23$$
$$d_2(u, v) = \|u - v\|_2 = \sqrt{4 + 64 + 49 + 36} = \sqrt{153}$$

7.55 Considere a função $f(t) = t^2 - 4t$ de $C[0, 3]$.

(a) Encontre $\|f\|_\infty$, (b) Esboce o gráfico de $f(t)$ no plano $\mathbf{R}^2$, (c) Encontre $\|f\|_1$, (d) Encontre $\|f\|_2$.

(a) Queremos calcular $\|f\|_\infty = \max(|f(t)|)$. Como $f(t)$ é derivável em [0, 3], $|f(t)|$ atinge um máximo num ponto crítico de $f(t)$ (ou seja, quando a derivada $f'(t) = 0$), ou num extremidade de [0, 3]. Como $f'(t) = 2t - 4$, tomamos $2t - 4 = 0$ e obtemos $t = 2$ como ponto crítico. Calculando,

$$f(2) = 4 - 8 = -4, \qquad f(0) = 0 - 0 = 0, \qquad f(3) = 9 - 12 = -3$$

Assim, $\|f\|_\infty = |f(2)| = |-4| = 4$.

(b) Calculamos $f(t)$ para alguns valores de t em [0, 3], por exemplo,

t	0	1	2	3
$f(t)$	0	−3	−4	−3

Esboçamos os pontos de $\mathbf{R}^2$ e traçamos uma curva contínua pelos pontos, conforme Figura 7-8.

(c) Queremos calcular $\|f\|_1 = \int_0^3 |f(t)|\, dt$. Conforme indicado na Figura 7-8, $f(t)$ é negativa em [0, 3], portanto, $|f(t)| = -(t^2 - 4t) = 4t - t^2$.

Assim, $\|f\|_1 = \displaystyle\int_0^3 (4t - t^2)\, dt = \left(2t^2 - \frac{t^3}{3}\right)\Big|_0^3 = 18 - 9 = 9$

(d) $\|f\|_2^2 = \displaystyle\int_0^3 f(t)^2\, dt = \int_0^3 (t^4 - 8t^3 + 16t^2)\, dt = \left(\frac{t^5}{5} - 2t^4 + \frac{16t^3}{3}\right)\Big|_0^3 = \frac{153}{5}$.

Assim, $\|f\|_2 = \sqrt{\dfrac{153}{5}}$.

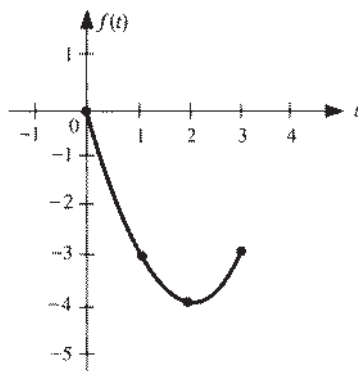



Figura 7-8

7.56 Demonstre o Teorema 7.24. Seja V um espaço vetorial normado. Então a função $d(u, v) = \|u - v\|$ satisfaz os três axiomas de espaço métrico, como segue.

$[M_1]$ $d(u, v) \geq 0$ e $d(u, v) = 0$ se, e só se, $u = v$.
$[M_2]$ $d(u, v) = d(v, u)$.
$[M_3]$ $d(u, v) \leq d(u, w) + d(w, v)$.

Se $u \neq v$, então $u - v \neq 0$ e, portanto, $d(u, v) = \|u - v\| > 0$. Também $d(u, u) = \|u - u\| = \|0\| = 0$. Assim, $[M_1]$ é válido. Também temos

$$d(u, v) = \|u - v\| = \|-1(v - u)\| = |-1|\|v - u\| = \|v - u\| = d(v, u)$$

e $$d(u, v) = \|u - v\| = \|(u - w) + (w - v)\| \leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v)$$

Assim, $[M_2]$ e $[M_3]$ são válidos.

Problemas Complementares

Produtos internos

7.57 Verifique que define um produto interno de $\mathbf{R}^2$, com $u = (x_1, x_2)$ e $v = (y_1, y_2)$, a função

$$f(u, v) = x_1 y_1 - 2x_1 y_2 - 2x_2 y_1 + 5x_2 y_2$$

7.58 Encontre os valores de k que fazem um produto interno de $\mathbf{R}^2$, com $u = (x_1, x_2)$ e $v = (y_1, y_2)$, da função

$$f(u, v) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + kx_2 y_2$$

7.59 Considere os vetores $u = (1, -3)$ e $v = (2, 5)$ de $\mathbf{R}^2$. Encontre

(a) $\langle u, v \rangle$ em relação ao produto interno canônico de $\mathbf{R}^2$.

(b) $\langle u, v \rangle$ em relação ao produto interno de $\mathbf{R}^2$ do Problema 7.57.

(c) $\|v\|$ usando o produto interno canônico de $\mathbf{R}^2$.

(d) $|v\|$ usando o produto interno de $\mathbf{R}^2$ do Problema 7.57.

7.60 Mostre que cada uma das funções seguintes não é um produto interno de $\mathbf{R}^3$, com $u = (x_1, x_2, x_3)$ e $v = (y_1, y_2, y_3)$.

(a) $\langle u, v \rangle = x_1 y_1 + x_2 y_2$, (b) $\langle u, v \rangle = x_1 y_2 x_3 + y_1 x_2 y_3$.

7.61 Seja V o espaço vetorial das matrizes $m \times n$ sobre $\mathbf{R}$. Mostre que $\langle A, B \rangle = \operatorname{tr}(B^T A)$ define um produto interno de V.

7.62 Suponha que $|\langle u, v \rangle| = \|u\| \|v\|$. (Observe que isso é a "igualdade" de Cauchy-Schwarz.) Mostre que u e v são linearmente dependentes.

7.63 Sejam $f(u, v)$ e $g(u, v)$ produtos internos de um espaço vetorial V sobre $\mathbf{R}$. Prove.

(a) A soma $f + g$ é um produto interno de V, sendo $(f + g)(u, v) = f(u, v) + g(u, v)$.

(b) O múltiplo kf, com $k > 0$, é um produto interno de V, sendo $(kf)(u, v) = kf(u, v)$.

Ortogonalidade, complementos ortogonais, conjuntos ortogonais

7.64 Seja V o espaço vetorial dos polinômios sobre $\mathbf{R}$ de grau ≤ 2 com o produto interno definido por $\langle f, g \rangle = \int_0^1 f(t) g(t)\, dt$. Encontre uma base do subespaço W ortogonal a $h(t) = 2t + 1$.

7.65 Encontre uma base do subespaço W de $\mathbf{R}^4$ ortogonal a $u_1 = (1, -2, 3, 4)$ e $u_2 = (3, -5, 7, 8)$.

7.66 Encontre uma base do subespaço W de $\mathbf{R}^5$ ortogonal aos vetores $u_1 = (1, 1, 3, 4, 1)$ e $u_2 = (1, 2, 1, 2, 1)$.

7.67 Seja $w = (1, -2, -1, 3)$ um vetor de $\mathbf{R}^4$. Encontre

(a) uma base ortogonal de $w^\perp$, (b) uma base ortonormal de $w^\perp$.

7.68 Seja W o subespaço de $\mathbf{R}^4$ ortogonal a $u_1 = (1, 1, 2, 2)$ e $u_2 = (0, 1, 2, -1)$. Encontre

(a) uma base ortogonal de W, (b) uma base ortonormal de W. (Compare com o Problema 7.65.)

7.69 Seja S o conjunto constituído dos vetores de $\mathbf{R}^4$ dados.

$$u_1 = (1, 1, 1, 1), \qquad u_2 = (1, 1, -1, -1), \qquad u_3 = (1, -1, 1, -1), \qquad u_4 = (1, -1, -1, 1)$$

(a) Mostre que S é ortogonal e uma base de $\mathbf{R}^4$.

(b) Escreva $v = (1, 3, -5, 6)$ como uma combinação linear de u_1, u_2, u_3, u_4.

(c) Encontre as coordenadas de um vetor $v = (a, b, c, d)$ arbitrário de $\mathbf{R}^4$ em relação à base S.

(d) Normalize S para obter um conjunto ortonormal de $\mathbf{R}^4$.

7.70 Considere $\mathbf{M} = \mathbf{M}_{2,2}$ com o produto interno $\langle A, B\rangle = \text{tr}(B^T A)$. Mostre que é uma base ortonormal de **M** o conjunto

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

7.71 Considere $\mathbf{M} = \mathbf{M}_{2,2}$ com o produto interno $\langle A, B\rangle = \text{tr}(B^T A)$. Encontre uma base para o complemento ortogonal (a) das matrizes diagonais, (b) das matrizes simétricas.

7.72 Seja $\{u_1, u_2, ..., u_r\}$ um conjunto ortogonal de vetores. Mostre que, dados escalares $a_1, a_2, ..., a_r$ quaisquer, o conjunto $\{a_1u_1, a_2u_2, ..., a_ru_r\}$ é ortogonal.

7.73 Sejam U e W subespaços de um espaço com produto interno V de dimensão finita. Mostre que

(a) $(U + W)^\perp = U^\perp \cap W^\perp$, (b) $(U \cap W)^\perp = U^\perp + W^\perp$.

Projeções, algoritmo de Gram-Schmidt, aplicações

7.74 Encontre o coeficiente de Fourier c e a projeção cw de v sobre w, com

(a) $v = (2, 3, -5)$ e $w = (1, -5, 2)$ em $\mathbf{R}^3$.

(b) $v = (1, 3, 1, 2)$ e $w = (1, -2, 7, 4)$ em $\mathbf{R}^4$.

(c) $v = t^2$ e $w = t + 3$ em $\mathbf{P}(t)$, com produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$.

(d) $v = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $w = \begin{bmatrix} 1 & 1 \\ 5 & 5 \end{bmatrix}$ em $\mathbf{M} = \mathbf{M}_{2,2}$, com produto interno $\langle A, B\rangle = \text{tr}(B^T A)$.

7.75 Seja U o subespaço de $\mathbf{R}^4$ gerado por

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, -1, 2, 2), \qquad v_3 = (1, 2, -3, -4)$$

(a) Aplique o algoritmo de Gram-Schmidt para encontrar uma base ortogonal e uma base ortonormal de U.

(b) Encontre a projeção de $v = (1, 2, -3, 4)$ sobre U.

7.76 Seja $v = (1, 2, 3, 4, 6)$. Encontre a projeção de v sobre W ou, em outras palavras, encontre $w \in W$ que minimiza $\|v - w\|$, sendo W o subespaço de $\mathbf{R}^5$ gerado por

(a) $u_1 = (1, 2, 1, 2, 1)$ e $u_2 = (1, -1, 2, -1, 1)$, (b) $v_1 = (1, 2, 1, 2, 1)$ e $v_2 = (1, 0, 1, 5, -1)$.

7.77 Considere o subespaço $W = \mathbf{P}_2(t)$ de $\mathbf{P}(t)$ com o produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Encontre a projeção de $f(t) = t^3$ sobre W. (*Sugestão*: use os polinômios ortogonais $1, 2t - 1, 6t^2 - 6t + 1$ obtidos no Problema 7.22.)

7.78 Considere $\mathbf{P}(t)$ com o produto interno $\langle f, g\rangle = \int_{-1}^1 f(t)g(t)$ e o subespaço $W = \mathbf{P}_3(t)$.

(a) Encontre uma base ortogonal de W aplicando o algoritmo de Gram-Schmidt em $\{1, t, t^2, t^3\}$.

(b) Encontre a projeção de $f(t) = t^5$ sobre W.

Matrizes ortogonais

7.79 Encontre o número de todas as matrizes 2×2 ortogonais da forma $\begin{bmatrix} \frac{1}{3} & x \\ y & z \end{bmatrix}$ e exiba-as.

7.80 Encontre uma matriz 3×3 ortogonal P cujas duas primeiras linhas sejam múltiplos de $u = (1, 1, 1)$ e $v = (1, -2, 3)$, respectivamente.

7.81 Encontre uma matriz ortogonal simétrica P cuja primeira linha seja $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$. (Compare com o Problema 7.32.)

7.82 Dizemos que duas matrizes reais A e B são *ortogonalmente equivalentes* se existir uma matriz ortogonal P tal que $B = P^T AP$. Mostre que essa relação é de equivalência.

Produtos internos e matrizes positivas

7.83 Encontre a matriz A que representa o produto interno canônico de $\mathbf{R}^2$ em relação a cada uma das bases seguintes.

(a) $\{v_1 = (1, 4),\ \ v_2 = (2, -3)\}$, (b) $\{w_1 = (1, -3),\ \ w_2 = (6, 2)\}$.

7.84 Considere o produto interno de $\mathbf{R}^2$ dado por

$$f(u, v) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2, \qquad \text{com} \qquad u = (x_1, x_2) \qquad v = (y_1, y_2)$$

Encontre a matriz B que representa esse produto interno de $\mathbf{R}^2$ em relação a cada base do Problema 7.83.

7.85 Encontre a matriz C que representa o produto interno canônico de $\mathbf{R}^3$ em relação à base S formada pelos vetores $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 1)$, $u_3 = (1, -1, 3)$.

7.86 Seja $V = \mathbf{P}_2(t)$ com produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$.

(a) Encontre $\langle f, g\rangle$, com $f(t) = t + 2$ e $g(t) = t^2 - 3t + 4$.

(b) Encontre a matriz A do produto interno em relação à base $\{1, t, t^2\}$ of V.

(c) Verifique o Teorema 7.16, mostrando que $\langle f, g\rangle = [f]^T A[g]$ em relação à base $\{1, t, t^2\}$.

7.87 Decida quais das matrizes seguintes são positivas.

(a) $\begin{bmatrix} 1 & 3 \\ 3 & 5 \end{bmatrix}$, (b) $\begin{bmatrix} 3 & 4 \\ 4 & 7 \end{bmatrix}$, (c) $\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} 6 & -7 \\ -7 & 9 \end{bmatrix}$.

7.88 Sejam A e B matrizes positivas. Mostre que

(a) $A + B$ é positiva, (b) kA é positiva, com $k > 0$.

7.89 Seja B uma matriz real não singular. Mostre que (a) B^TB é simétrica, (b) B^TB é positiva.

Espaços com produto interno complexos

7.90 Verifique que

$$\langle a_1u_1 + a_2u_2 \quad b_1v_1 + b_2v_2\rangle = a_1\bar{b}_1\langle u_1, v_1\rangle + a_1\bar{b}_2\langle u_1, v_2\rangle + a_2\bar{b}_1\langle u_2, v_1\rangle + a_2\bar{b}_2\langle u_2, v_2\rangle$$

Mais geralmente, prove que $\langle \sum_{i=1}^m a_iu_i,\ \ \sum_{j=1}^n b_jv_j\rangle = \sum_{i,j} a_i\bar{b}_j\langle u_i, v_i\rangle$.

7.91 Considere $u = (1 + i,\ \ 3,\ \ 4 - i)$ e $v = (3 - 4i,\ \ 1 + i,\ \ 2i)$ de $\mathbf{C}^3$. Encontre

(a) $\langle u, v\rangle$, (b) $\langle v, u\rangle$, (c) $\|u\|$, (d) $\|v\|$, (e) $d(u, v)$.

7.92 Encontre o coeficiente de Fourier c e a projeção cw de

(a) $u = (3 + i,\ \ 5 - 2i)$ sobre $w = (5 + i,\ \ 1 + i)$ em $\mathbf{C}^2$.

(b) $u = (1 - i,\ \ 3i,\ \ 1 + i)$ sobre $w = (1,\ \ 2 - i,\ \ 3 + 2i)$ em $\mathbf{C}^3$.

7.93 Sejam $u = (z_1, z_2)$ e $v = (w_1, w_2)$ de $\mathbf{C}^2$. Verifique que a função dada é um produto interno de $\mathbf{C}^2$.

$$f(u, v) = z_1\bar{w}_1 + (1 + i)z_1\bar{w}_2 + (1 - i)z_2\bar{w}_1 + 3z_2\bar{w}_2$$

7.94 Encontre uma base ortogonal e uma base ortonormal do subespaço W de $\mathbf{C}^3$ gerado por $u_1 = (1, i, 1)$ e $u_2 = (1 + i,\ \ 0,\ \ 2)$.

7.95 Sejam $u = (z_1, z_2)$ e $v = (w_1, w_2)$ de $\mathbf{C}^2$. Encontre os valores de $a, b, c, d \in \mathbf{C}$ para os quais a função dada é um produto interno de $\mathbf{C}^2$.

$$f(u, v) = az_1\bar{w}_1 + bz_1\bar{w}_2 + cz_2\bar{w}_1 + dz_2\bar{w}_2$$

7.96 Demonstre a expansão seguinte de um produto interno de um espaço complexo V.

$$\langle u, v\rangle = \tfrac{1}{4}\|u + v\|^2 - \tfrac{1}{4}\|u - v\|^2 + \tfrac{1}{4}\|u + iv\|^2 - \tfrac{1}{4}\|u - iv\|^2$$

[Compare com o Problema 7.7(b).]

7.97 Seja V um espaço com produto interno real. Mostre que

(a) $\|u\| = \|v\|$ se, e só se, $\langle u+v, \ u-v\rangle = 0$.

(b) $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ se, e só se, $\langle u, v\rangle = 0$.

Dê contraexemplos que mostrem que as afirmações dadas não valem em, digamos, $\mathbf{C}^2$.

7.98 Encontre a matriz P que representa o produto interno canônico de $\mathbf{C}^3$ em relação à base $\{1, \ 1+i, \ 1-2i\}$.

7.99 Uma matriz complexa A é *unitária* se for invertível e $A^{-1} = A^H$. Alternativamente, A é unitária se suas linhas (colunas) formarem um conjunto ortonormal de vetores (em relação ao produto interno canônico de $\mathbf{C}^n$). Encontre uma matriz unitária cuja primeira linha seja (a) um múltiplo de $(1, \ 1-i)$, (b) um múltiplo de $(\frac{1}{2}, \ \frac{1}{2}i, \ \frac{1}{2}-\frac{1}{2}i)$.

Espaços vetoriais normados

7.100 Considere os vetores $u = (1, -3, 4, 1, -2)$ e $v = (3, 1, -2, -3, 1)$ de $\mathbf{R}^5$. Encontre

(a) $\|u\|_\infty$ e $\|v\|_\infty$, (b) $\|u\|_1$ e $\|v\|_1$, (c) $\|u\|_2$ e $\|v\|_2$, (d) $d_\infty(u, v), d_1(u, v), d_2(u, v)$

7.101 Repita o Problema 7.100 com $u = (1+i, \ 2-4i)$ e $v = (1-i, \ 2+3i)$ de $\mathbf{C}^2$.

7.102 Considere as funções $f(t) = 5t - t^2$ e $g(t) = 3t - t^2$ de $C[0, 4]$. Encontre

(a) $d_\infty(f, g)$, (b) $d_1(f, g)$, (c) $d_2(f, g)$

7.103 Demonstre que (a) $\|\cdot\|_1$ é uma norma de $\mathbf{R}^n$, (b) $\|\cdot\|_\infty$ é uma norma de $\mathbf{R}^n$.

7.104 Demonstre que (a) $\|\cdot\|_1$ é uma norma de $C[a, b]$, (b) $\|\cdot\|_\infty$ é uma norma de $C[a, b]$.

Respostas dos Problemas Complementares

Notação: $M = [R_1; \ R_2; \ \ldots]$ denota uma matriz M de linhas R_1, R_2, ... Também observe que as bases não são únicas.

7.58 $k > 9$

7.59 (a) -13, (b) -71, (c) $\sqrt{29}$, (d) $\sqrt{89}$

7.60 Seja $u = (0, 0, 1)$; então $\langle u, u\rangle = 0$ em ambos casos.

7.64 $\{7t^2 - 5t, \ 12t^2 - 5\}$

7.65 $\{(1, 2, 1, 0), \ (4, 4, 0, 1)\}$

7.66 $(-1, 0, 0, 0, 1), (-6, 2, 0, 1, 0), (-5, 2, 1, 0, 0)$

7.67 (a) $u_1 = (0, 0, 3, 1), u_2 = (0, 5, -1, 3), u_3 = (-14, -2, -1, 3)$,
(b) $u_1/\sqrt{10}, u_2/\sqrt{35}, u_3/\sqrt{210}$

7.68 (a) $(0, 2, -1, 0), (-15, 1, 2, 5)$, (b) $(0, 2, -1, 0)/\sqrt{5}, (-15, 1, 2, 5)/\sqrt{255}$

7.69 (b) $v = \frac{1}{4}(5u_1 + 3u_2 - 13u_3 + 9u_4)$,
(c) $[v] = \frac{1}{4}[a+b+c+d, \ a+b-c-d, \ a-b+c-d, \ a-b-c+d]$

7.71 (a) $[0, 1; \ 0, 0]$, $[0, 0; \ 1, 0]$, (b) $[0, -1; \ 1, 0]$

7.74 (a) $c = -\frac{23}{30}$, (b) $c = \frac{1}{7}$, (c) $c = \frac{15}{148}$, (d) $c = \frac{19}{26}$

7.75 (a) $w_1 = (1,1,1,1), w_2 = (0,-2,1,1), w_3 = (12,-4,-1,-7)$,
(b) $\text{proj}(v,U) = \frac{1}{5}(-1,12,3,6)$

7.76 (a) $\text{proj}(v,W) = \frac{1}{8}(23,25,30,25,23)$, (b) Primeiro, encontre uma base ortogonal de W, digamos, $w_1 = (1,2,1,2,1)$ e $w_2 = (0,2,0,-3,2)$. Então, $\text{proj}(v,W) = \frac{1}{17}(34,76,34,56,42)$.

7.77 $\text{proj}(f,W) = \frac{3}{2}t^2 - \frac{3}{5}t + \frac{1}{20}$

7.78 (a) $\{1,\ t,\ 3t^2-1,\ 5t^3-3t\}$, (b) $\text{proj}(f,W) = \frac{10}{9}t^3 - \frac{5}{21}t$

7.79 Quatro: $[a,b;\quad b,-a]$, $[a,b;\quad -b,-a]$, $[a,-b;\quad b,a]$, $[a,-b;\quad -b,-a]$, where $a = \frac{1}{3}$ e $b = \frac{1}{3}\sqrt{8}$

7.80 $P = [1/a, 1/a, 1/a;\quad 1/b, -2/b, 3/b;\quad 5/c, -2/c, -3/c]$, com $a = \sqrt{3}, b = \sqrt{14}, c = \sqrt{38}$

7.81 $\frac{1}{3}[1,2,2;\quad 2,-2,1;\quad 2,1,-2]$

7.83 (a) $[17,-10;\quad -10,13]$, (b) $[10,0;\quad 0,40]$

7.84 (a) $[65,-68;\quad -68,73]$, (b) $[58,8;\quad 8,8]$

7.85 $[3,4,3;\quad 4,6,2;\quad 3,2,11]$

7.86 (a) $\frac{83}{12}$, (b) $[1,a,b;\quad a,b,c;\quad b,c,d]$, com $a = \frac{1}{2}$, $b = \frac{1}{3}$, $c = \frac{1}{4}$, $d = \frac{1}{5}$

7.87 (a) Não é, (b) É, (c) Não é, (d) É.

7.91 (a) $-4i$, (b) $4i$, (c) $\sqrt{28}$, (d) $\sqrt{31}$, (e) $\sqrt{59}$

7.92 (a) $c = \frac{1}{28}(19-5i)$, (b) $c = \frac{1}{19}(3+6i)$

7.94 $\{v_1 = (1,i,1)/\sqrt{3},\quad v_2 = (2i,\ 1-3i,\ 3-i)/\sqrt{24}\}$

7.95 a e d reais e positivos, $c = \bar{b}$ e $ad - bc$ positivos.

7.97 $u = (1,2),\quad v = (i,2i)$

7.98 $P = [1,\ 1-i,\ 1+2i;\quad 1+i,\ 2,\ -1+3i;\quad 1-2i,\ -1-3i,\ 5]$

7.99 (a) $(1/\sqrt{3})[1,\ 1-i;\quad 1+i,\ -1]$,
(b) $[a,\ ai,\ a-ai;\quad bi,\ b,\ 0;\quad a,\ ai,\ -a-ai]$, com $a = \frac{1}{2}$ e $b = 1/\sqrt{2}$.

7.100 (a) 4 e 3, (b) 11 e 10, (c) $\sqrt{31}$ e $\sqrt{24}$, (d) 6, 19, 9

7.101 (a) $\sqrt{20}$ e $\sqrt{13}$, (b) $\sqrt{2}+\sqrt{20}$ e $\sqrt{2}+\sqrt{13}$, (c) $\sqrt{22}$ e $\sqrt{15}$, (d) $7, 9, \sqrt{53}$

7.102 (a) 8, (b) 16, (c) $16/\sqrt{3}$.

Capítulo 8

Determinantes

8.1 INTRODUÇÃO

A cada matriz quadrada $A = [a_{ij}]$ de ordem n está associado um escalar especial, denominado *determinante* de A e denotado por $\det(A)$, ou $|A|$ ou

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Enfatizamos que uma tabela $n \times n$ de escalares emoldurada por segmentos de reta, denominada *determinante de ordem n*, não é uma matriz e somente denota o determinante da tabela de escalares (ou seja, da matriz emoldurada).

A função determinante foi descoberta durante a investigação de sistemas de equações lineares. Veremos que o determinante é uma ferramenta indispensável na investigação e na obtenção de propriedades de matrizes quadradas.

A definição do determinante e a maioria de suas propriedades também se aplicam ao caso em que as entradas da matriz provêm de um anel comutativo.

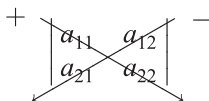
Começamos com o caso especial de determinantes de ordens 1, 2 e 3. Depois definimos um determinante de ordem arbitrária. Essa definição geral é precedida por uma discussão de permutações necessária para a nossa definição geral de determinante.

8.2 DETERMINANTES DE ORDENS 1 E 2

Os determinantes de ordens 1 e 2 são definidos como segue.

$$|a_{11}| = a_{11} \qquad \text{e} \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Assim, o determinante de uma matriz $A = [a_{11}]$ de ordem 1 é o próprio escalar a_{11}, isto é, $\det(A) = |a_{11}| = a_{11}$. O determinante de ordem dois pode ser facilmente lembrado usando o diagrama a seguir.

$$\begin{matrix} + & & & - \\ & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & & \end{matrix}$$

Ou seja, o determinante é igual ao produto dos elementos ao longo da seta marcada com o sinal de soma menos o produto dos elementos ao longo da seta marcada com o sinal de diferença. (Existe um diagrama análogo para determinantes de ordem 3, mas não para determinantes de ordens superiores.)

Exemplo 8.1

(a) Como o determinante de ordem 1 é o próprio escalar, temos

$$\det(27) = 27, \qquad \det(-7) = -7, \qquad \det(t-3) = t-3$$

(b) $\begin{vmatrix} 5 & 3 \\ 4 & 6 \end{vmatrix} = 5(6) - 3(4) = 30 - 12 = 18, \qquad \begin{vmatrix} 3 & 2 \\ -5 & 7 \end{vmatrix} = 21 + 10 = 31$

Aplicação a equações lineares

Considere duas equações lineares com duas incógnitas, digamos,

$$\begin{aligned} a_1 z + b_1 y &= c_1 \\ a_2 x + b_2 y &= c_2 \end{aligned}$$

Seja $D = a_1 b_2 - a_2 b_1$ o determinante da matriz de coeficientes. Então o sistema tem uma única solução se, e só se, $D \neq 0$. Nesse caso, a única solução pode ser completamente expressa em termos de determinantes, como segue.

$$x = \frac{N_x}{D} = \frac{b_2 c_1 - b_1 c_2}{a_1 b_2 - a_2 b_1} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \qquad y = \frac{N_y}{D} = \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

Observe que D aparece no denominador de ambos os quocientes. Os numeradores N_x e N_y dos quocientes que determinam x e y, respectivamente, podem ser obtidos substituindo a coluna dos termos constantes no lugar da coluna dos coeficientes da incógnita dada na matriz de coeficientes. Por outro lado, se $D = 0$, então o sistema pode não ter solução ou, então, ter mais de uma solução.

Exemplo 8.2 Resolva com determinantes o sistema $\begin{cases} 4x - 3y = 15 \\ 2x + 5y = \ \ 1 \end{cases}$

Primeiro, calculamos o determinante D da matriz de coeficientes.

$$D = \begin{vmatrix} 4 & -3 \\ 2 & 5 \end{vmatrix} = 4(5) - (-3)(2) = 20 + 6 = 26$$

Como $D \neq 0$, o sistema possui solução única. Para obter os numeradores N_x e N_y, simplesmente substituímos, na matriz de coeficientes, os coeficientes de x e y, respectivamente, pelos termos constantes e, então, tomamos seus determinantes.

$$N_x = \begin{vmatrix} 15 & -3 \\ 1 & 5 \end{vmatrix} = 75 + 3 = 78 \qquad N_y = \begin{vmatrix} 4 & 15 \\ 2 & 1 \end{vmatrix} = 4 - 30 = -26$$

Então, a única solução do sistema é

$$x = \frac{N_x}{D} = \frac{78}{26} = 3, \qquad y = \frac{N_y}{D} = \frac{-26}{26} = -1$$

8.3 DETERMINANTES DE ORDEM 3

Considere uma matriz qualquer $A = [a_{ij}]$ de ordem 3. O determinante de A é definido como segue.

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Observe que há seis produtos, cada um consistindo de três elementos da matriz original. Três desses produtos recebem o sinal positivo (mantêm seu sinal) e os outros três recebem o sinal negativo (mudam seu sinal).

Os diagramas da Figura 8-1 podem ajudar a lembrar desses seis produtos de det(A). Ou seja, o determinante é igual à soma dos produtos dos elementos ao longo das três setas assinaladas com o sinal de soma na Figura 8-1, mais a soma dos negativos dos produtos dos elementos que estão nas setas assinaladas com o sinal de diferença. Enfatizamos que não existem recursos visuais com os quais seja possível lembrar determinantes de ordens superiores.

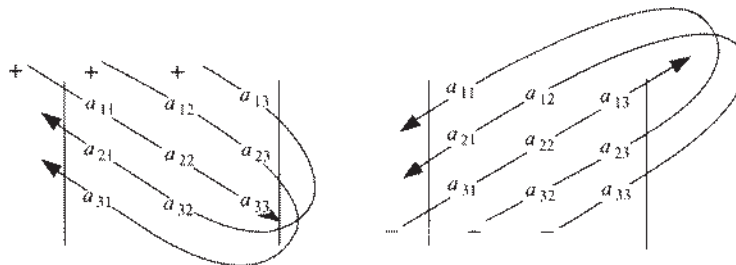



Figura 8-1

Exemplo 8.3 Sejam $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 2 & 1 \\ -4 & 5 & -1 \\ 2 & -3 & 4 \end{bmatrix}$. Encontre det($A$) e det($B$).

Usamos os diagramas da Figura 8-1.

$$\begin{aligned} \det(A) &= 2(5)(4) + 1(-2)(1) + 1(-3)(0) - 1(5)(1) - (-3)(-2)(2) - 4(1)(0) \\ &= 40 - 2 + 0 - 5 - 12 - 0 = 21 \\ \det(B) &= 60 - 4 + 12 - 10 - 9 + 32 = 81 \end{aligned}$$

Forma alternativa para o determinante de ordem 3

O determinante de uma matriz $A = [a_{ij}]$ de ordem 3 pode ser reescrito como segue.

$$\begin{aligned} \det(A) &= a_{11}(a_{22}a_{23} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

que é uma combinação linear de três determinantes de ordem 2 cujos coeficientes (com sinais alternados) formam a primeira linha da matriz dada. Essa combinação linear pode ser indicada como segue.

$$a_{11}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Observe que cada matriz 2×2 pode ser obtida suprimindo, na matriz original, a linha e a coluna que contêm seu coeficiente.

Exemplo 8.4

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} &= 1\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} \\ &= 1\begin{vmatrix} -2 & 3 \\ 5 & -1 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 \\ 0 & -1 \end{vmatrix} + 3\begin{vmatrix} 4 & -2 \\ 0 & 5 \end{vmatrix} \\ &= 1(2 - 15) - 2(-4 + 0) + 3(20 + 0) = -13 + 8 + 60 = 55 \end{aligned}$$

8.4 PERMUTAÇÕES

Uma permutação σ do conjunto {1, 2, ..., n} é uma bijeção desse conjunto em si mesmo ou, equivalentemente, uma reordenação dos números 1, 2, ..., n. Uma permutação σ é denotada por

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} \qquad \text{ou} \qquad \sigma = j_1 j_2 \cdots j_n, \qquad \text{com } j_i = \sigma(i)$$

O conjunto de todas essas permutações é denotado por S_n e o número dessas permutações é $n!$. Se $\sigma \in S_n$, então a aplicação inversa $\sigma^{-1} \in S_n$ e, se $\sigma, \tau \in S_n$, então a aplicação composta $\sigma \circ \tau \in S_n$. Também a transformação identidade $\varepsilon = \sigma \circ \sigma^{-1} \in S_n$. (Na verdade, $\varepsilon = 123 \ldots n$.)

Exemplo 8.5

(a) Existem $2! = 2 \cdot 1 = 2$ permutações em S_2, a saber, 12 e 21.
(b) Existem $3! = 3 \cdot 2 \cdot 1 = 6$ permutações em S_3, a saber, 123, 132, 213, 231, 312 e 321.

Sinal (paridade) de uma permutação

Considere uma permutação arbitrária σ de S_n, digamos, $\sigma = j_1 j_2 \cdots j_n$. Dizemos que σ é uma permutação par ou ímpar se existir um número par ou ímpar de inversões em σ. Por uma *inversão* em σ entendemos um par de inteiros (i, k) tais que $i > k$, mas com i precedendo k em σ. Então definimos o sinal ou a paridade de σ, e escrevemos sgn σ, por

$$\text{sgn } \sigma = \begin{cases} 1 & \text{se } \sigma \text{ é par} \\ -1 & \text{se } \sigma \text{ é ímpar} \end{cases}$$

Exemplo 8.6

(a) Encontre o sinal de $\sigma = 35142$ em S_5.

Para cada elemento k, contamos o número de elementos i tais que $i > k$, mas com i precedendo k em σ. Temos

2 números (3 e 5) maiores do que e precedendo 1,
3 números (3, 5 e 4) maiores do que e precedendo 2,
1 número (5) maior do que e precedendo 4.

(Não existem números maiores do que e precedendo 3 e 5.) Como há um total de seis inversões, σ é par e sgn $\sigma = 1$.

(b) A permutação identidade $\varepsilon = 123 \ldots n$ é par, pois não há inversões em ε.
(c) Em S_2, a permutação 12 é par e 21 é ímpar. Em S_3, as permutações 123, 231, 312 são pares e as permutações 132, 213, 321 são ímpares.
(d) Seja τ a permutação que troca de posição os dois números i e j, deixando os demais fixos. Ou seja,

$$\tau(i) = j, \qquad \tau(j) = i, \qquad \tau(k) = k, \qquad \text{com} \quad k \neq i, j$$

Dizemos que τ é uma *transposição*. Se $i < j$, então existem $2(j - i) - 1$ inversões em τ e, portanto, a transposição τ é ímpar.

OBSERVAÇÃO Podemos mostrar que, dado qualquer n, a metade das permutações de S_n são pares e a outra metade são ímpares. Por exemplo, 3 das 6 permutações de S_3 são pares e 3 são ímpares.

8.5 DETERMINANTES DE ORDEM ARBITRÁRIA

Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n sobre um corpo K.

Considere um produto de n elementos de A tal que um, e apenas um, desses elementos vêm de cada linha e um, e apenas um, vêm de cada coluna. Um tal produto pode ser escrito na forma

$$a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

em que os fatores pertencem a linhas sucessivas, sendo os primeiros de seus dois índices na ordem natural 1, 2, ..., n. Agora, como os fatores vêm de colunas diferentes, a sequência dos segundos de seus dois índices forma uma permutação $\sigma = j_1 j_2 \cdots j_n$ de S_n. Reciprocamente, cada permutação de S_n determina um produto desse tipo. Assim, a matriz A fornece $n!$ desses produtos.

DEFINIÇÃO O determinante de $A = [a_{ij}]$, denotado por det(A) ou $|A|$, é a soma de todos os $n!$ produtos descritos acima, sendo cada um desses produtos multiplicado por sgn σ, isto é,

$$|A| = \sum_{\sigma} (\text{sgn } \sigma) a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

ou

$$|A| = \sum_{\sigma \in S_n} (\text{sgn } \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

Dizemos que a matriz quadrada A de ordem n tem um determinante de ordem n.

O exemplo seguinte mostra que a definição dada coincide com as definições precedentes de determinantes de ordem 1, 2 e 3.

Exemplo 8.7

(a) Seja $A = [a_{11}]$ uma matriz 1×1. Como S_1 só tem uma permutação, que é par, $\det(A) = a_{11}$, o próprio elemento.

(b) Seja $A = [a_{ij}]$ uma matriz 2×2. Em S_2 a permutação 12 é par e a permutação 21 é ímpar. Logo,

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

(c) Seja $A = [a_{ij}]$ uma matriz 3×3. Em S_3 as permutações 123, 231, 312 são pares e as permutações 321, 213, 132 são ímpares. Logo,

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

OBSERVAÇÃO Quando n cresce, o número de termos no determinante se torna astronômico. Por isso, usamos métodos indiretos para calcular determinantes em vez da definição de determinante. Na verdade, demonstramos várias propriedades dos determinantes que nos permitem simplificar consideravelmente as contas. Em particular, mostramos que um determinante de ordem n é igual a uma combinação linear de determinantes de ordem $n - 1$, como no caso $n = 3$ já visto.

8.6 PROPRIEDADES DE DETERMINANTES

Vejamos as propriedades básicas do determinante.

Teorema 8.1 O determinante de uma matriz A coincide com o de sua transposta A^T, isto é, $|A| = |A^T|$.

Com esse teorema (demonstrado no Problema 8.22), vemos que qualquer teorema sobre o determinante de uma matriz A relacionado com as linhas de A tem um teorema análogo relacionado com as colunas de A.

O teorema seguinte (demonstrado no Problema 8.24) apresenta certos casos em que o determinante pode ser obtido imediatamente.

Teorema 8.2 Seja A uma matriz quadrada.

(i) Se A tem uma linha (coluna) de zeros, então $|A| = 0$.
(ii) Se A tem duas linhas (colunas) idênticas, então $|A| = 0$.

(iii) Se A é triangular, isto é, se A possui zeros acima ou abaixo da diagonal, então $|A|$ é igual ao produto dos elementos diagonais. Assim, em particular, $|I| = 1$, sendo I a matriz identidade.

O teorema seguinte (demonstrado nos Problemas 8.23 e 8.25) mostra como o determinante de uma matriz é afetado pelas operações elementares com as linhas e as colunas.

Teorema 8.3 Suponha que B seja obtida a partir de A por uma operação elementar com as linhas (colunas).

(i) Se duas linhas (colunas) trocaram de posição, então $|B| = -|A|$.

(ii) Se uma linha (coluna) de A foi multiplicada por um escalar k, então $|B| = k|A|$.

(iii) Se um múltiplo de uma linha (coluna) de A foi somado a uma outra linha (coluna) de A, então $|B| = |A|$.

Propriedades principais dos determinantes

Enunciamos, agora, duas das mais importantes e úteis propriedades de determinantes.

Teorema 8.4 O determinante de um produto de duas matrizes A e B é o produto de seus determinantes, isto é,

$$\det(AB) = \det(A)\,\det(B)$$

Esse teorema afirma que o determinante é uma função multiplicativa.

Teorema 8.5 Seja A uma matriz quadrada. São equivalentes as afirmações dadas.

(i) A é invertível, isto é, A possui uma inversa A^{-1}.
(ii) $AX = 0$ possui apenas a solução nula.
(iii) O determinante de A é não nulo, isto é, $\det(A) \neq 0$.

OBSERVAÇÃO Dependendo do autor e do texto, uma matriz não singular A é definida como uma matriz invertível, ou como uma matriz A para a qual $|A| \neq 0$, ou como uma matriz A tal que $AX = 0$ só tem a solução nula. O teorema precedente mostra que todas essas definições são equivalentes.

Demonstramos os Teoremas 8.4 e 8.5 (nos Problemas 8.29 e 8.28, respectivamente) utilizando a teoria de matrizes elementares e o lema seguinte (demonstrado no Problema 8.26), que é um caso especial do Teorema 8.4.

Lema 8.6 Seja E uma matriz elementar. Então, dada qualquer matriz A, $|EA| = |E||A|$.

Vimos que duas matrizes A e B são semelhantes se existir uma matriz não singular P tal que $B = P^{-1}AP$. Usando a propriedade multiplicativa do determinante (Teorema 8.4), pode-se demonstrar facilmente (Problema 8.31) o teorema seguinte.

Teorema 8.7 Se A e B são matrizes semelhantes, então $|A| = |B|$.

8.7 MENORES E COFATORES

Considere uma matriz quadrada $A = [a_{ij}]$ de ordem n. Denotemos por M_{ij} a submatriz de A de ordem $(n-1)$ obtida suprimindo sua i-ésima linha e sua j-ésima coluna. O determinante $|M_{ij}|$ é denominado *menor* do elemento a_{ij} de A e o *cofator* de a_{ij}, denotado por A_{ij}, é o menor com "sinal", como segue.

$$A_{ij} = (-1)^{i+j}|M_{ij}|$$

Observe que os "sinais" $(-1)^{i+j}$ que acompanham os menores formam um padrão de tabuleiro de xadrez com o sinal da soma na diagonal.

$$\begin{bmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ \multicolumn{5}{c}{\dots\dots\dots\dots\dots\dots} \end{bmatrix}$$

Enfatizamos que M_{ij} denota uma matriz, enquanto A_{ij} denota um escalar.

OBSERVAÇÃO Frequentemente, o sinal $(-1)^{i+j}$ do cofator A_{ij} é obtido com o padrão do tabuleiro de xadrez. Mais precisamente, começamos com o sinal da soma e alternamos o sinal, ou seja,

$$+, -, +, -, \dots,$$

contando da diagonal principal para a casa apropriada do tabuleiro.

Exemplo 8.8 Seja $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Encontre os menores e os cofatores seguintes. (a) $|M_{23}|$ e A_{23}, (b) $|M_{31}|$ e A_{31}.

(a) $|M_{23}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = 8 - 14 = -6$, e, portanto, $A_{23} = (-1)^{2+3}|M_{23}| = -(-6) = 6$

(b) $|M_{31}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 12 - 15 = -3$, e, portanto, $A_{31} = (-1)^{1+3}|M_{31}| = +(-3) = -3$

Expansão de Laplace

Vale o teorema seguinte (demonstrado no Problema 8.32).

Teorema 8.8 (Laplace) O determinante de uma matriz quadrada $A = [a_{ij}]$ é igual à soma dos produtos obtidos multiplicando os elementos de qualquer linha (coluna) por seus respectivos cofatores.

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^{n} a_{ij}A_{ij}$$

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}$$

Essas fórmulas para $|A|$ são denominadas *expansões de Laplace* do determinante de A pela i-ésima linha e j-ésima coluna. Junto com as operações elementares com as linhas (colunas), oferecem um método para simplificar o cálculo de $|A|$, conforme veremos.

8.8 CÁLCULO DE DETERMINANTES

O algoritmo seguinte reduz o cálculo de um determinante de ordem n ao cálculo de um determinante de ordem $n - 1$.

Algoritmo 8.1 (Redução de ordem de um determinante) É dada uma matriz não nula $A = [a_{ij}]$ quadrada de ordem n, com $n > 1$.

Passo 1 Escolha um elemento $a_{ij} = 1$ ou, na falta desse, $a_{ij} \neq 0$.

Passo 2 Usando a_{ij} como pivô, use operações elementares com as linhas (colunas) para colocar zeros em todas as outras posições da coluna (linha) contendo a_{ij}.

Passo 3 Expanda o determinante pela coluna (linha) contendo a_{ij}.

Destacamos as observações seguintes.

OBSERVAÇÃO 1 O Algoritmo 8.1 costuma ser usado para determinantes de ordem 4, ou superior. Para determinantes de ordem menor do que 4, usamos as fórmulas específicas para o determinante.

OBSERVAÇÃO 2 A eliminação gaussiana ou, equivalentemente, o uso repetido do Algoritmo 8.1, junto com trocas de linhas, podem ser usados para transformar uma matriz A numa matriz triangular superior cujo determinante seja o produto de suas entradas diagonais. Contudo, devemos manter um registro do número de trocas de linhas, já que cada troca de linha troca o sinal do determinante.

Exemplo 8.9 Use o Algoritmo 8.1 para calcular o determinante de $A = \begin{bmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{bmatrix}$.

Usamos a_{23} como pivô e colocamos zeros nas outras posições da terceira coluna, isto é, aplicamos as operações elementares com as linhas "Substituir R_1 por $-2R_2 + R_1$", "Substituir R_3 por $3R_2 + R_3$" e "Substituir R_4 por $R_2 + R_4$". Pelo Teorema 8.3(iii), o valor do determinante não se altera com essas operações. Assim,

$$|A| = \begin{vmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix}$$

Agora, expandimos a terceira coluna. Mais precisamente, suprimimos todos os termos iguais a zero e usamos o fato de que o sinal do menor M_{23} é $(-1)^{2+3} = -1$. Assim,

$$|A| = -\begin{vmatrix} 1 & 2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & -2 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix} = -(4 - 18 + 5 - 30 - 3 + 4) = -(-38) = 38$$

8.9 ADJUNTA CLÁSSICA

Sejam $A = [a_{ij}]$ uma matriz $n \times n$ sobre um corpo K e A_{ij} o cofator de a_{ij}. A *adjunta clássica* de A, denotada por adj A, é a transposta da matriz de cofatores de A, a saber,

$$\text{adj } A = [A_{ij}]^T$$

Dizemos "adjunta clássica", em vez de simplesmente "adjunta", porque, hoje me dia, o termo "adjunta" é utilizado para um conceito totalmente diferente (ver Capítulo 13).

Exemplo 8.10 Seja $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$. Vejamos os cofatores dos nove elementos de A.

$$A_{11} = +\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -18, \quad A_{12} = -\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} = 2, \quad A_{13} = +\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} = 4$$

$$A_{21} = -\begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} = -11, \quad A_{22} = +\begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} = 14, \quad A_{23} = -\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 5$$

$$A_{31} = +\begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} = -10, \quad A_{32} = -\begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} = -4, \quad A_{33} = +\begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8$$

A transposta da matriz dos cofatores fornece a adjunta clássica de A, ou seja,

$$\text{adj } A = \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix}$$

Vale o teorema seguinte (demonstrado no Problema 8.34).

Teorema 8.9 Seja A uma matriz quadrada qualquer. Então

$$A(\text{adj } A) = (\text{adj } A)A = |A|I$$

onde I é a matriz identidade. Assim, se $|A| \neq 0$,

$$A^{-1} = \frac{1}{|A|}(\text{adj } A)$$

Exemplo 8.11 Seja A a matriz do Exemplo 8.10. Temos

$$\det(A) = -40 + 6 + 0 - 16 + 4 + 0 = -46$$

Assim, A possui uma inversa e, pelo Teorema 8.9,

$$A^{-1} = \frac{1}{|A|}(\text{adj } A) = -\frac{1}{46}\begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix} = \begin{bmatrix} \frac{9}{23} & \frac{11}{46} & \frac{5}{23} \\ -\frac{1}{23} & -\frac{7}{23} & \frac{2}{23} \\ -\frac{2}{23} & -\frac{5}{46} & \frac{4}{23} \end{bmatrix}$$

8.10 APLICAÇÕES A EQUAÇÕES LINEARES, REGRA DE CRAMER

Considere um sistema $AX = B$ de n equações lineares com n incógnitas. Seja $A = [a_{ij}]$ a matriz (quadrada) de coeficientes e $B = [b_i]$ o vetor coluna dos termos constantes. Seja A_i a matriz obtida de A substituindo a coluna i de A pelo vetor coluna B. Além disso, denotemos

$$D = \det(A), \qquad N_1 = \det(A_1), \qquad N_2 = \det(A_2), \qquad \ldots, \qquad N_n = \det(A_n)$$

A relação fundamental entre esses determinantes e a solução do sistema $AX = B$ é dada no teorema a seguir.

Teorema 8.10 O sistema (quadrado) $AX = B$ possui uma única solução se, e só se, $D \neq 0$. Nesse caso, a única solução é dada por

$$x_1 = \frac{N_1}{D}, \qquad x_2 = \frac{N_2}{D}, \qquad \ldots, \qquad x_n = \frac{N_n}{D}$$

Esse teorema (demonstrado no Problema 8.35) é conhecido como *Regra de Cramer* para a resolução de sistemas de equações lineares. Enfatizamos que o teorema apenas se refere a sistemas com o mesmo número de equações e incógnitas, e apenas fornece a solução no caso $D \neq 0$. De fato, se $D = 0$, o teorema não nos diz se o sistema possui solução, ou não. Contudo, no caso de um sistema homogêneo, temos o resultado útil seguinte (demonstrado no Problema 8.54).

Teorema 8.11 Um sistema homogêneo quadrado $AX = 0$ tem alguma solução não nula se, e só se, $D = |A| = 0$.

Exemplo 8.12 Usando determinantes, resolva o sistema $\begin{cases} x + \ y + \ z = \ 5 \\ x - 2y - 3z = -1 \\ 2x + \ y - \ z = \ 3 \end{cases}$

Em primeiro lugar, calculamos o determinante D da matriz de coeficientes.

$$D = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -2 & -3 \\ 2 & 1 & -1 \end{vmatrix} = 2 - 6 + 1 + 4 + 3 + 1 = 5$$

Como $D \neq 0$, o sistema tem solução única. Para calcular N_x, N_y, N_z substituímos, respectivamente, os coeficientes de x, y, z na matriz de coeficientes pelos termos constantes, obtendo

$$N_x = \begin{vmatrix} 5 & 1 & 1 \\ -1 & -2 & -3 \\ 3 & 1 & -1 \end{vmatrix} = 20, \qquad N_y = \begin{vmatrix} 1 & 5 & 1 \\ 1 & -1 & -3 \\ 2 & 3 & -1 \end{vmatrix} = -10, \qquad N_z = \begin{vmatrix} 1 & 1 & 5 \\ 1 & -2 & -1 \\ 2 & 1 & 3 \end{vmatrix} = 15$$

Assim, a solução única do sistema é $x = N_x/D = 4$, $y = N_y/D = -2$, $z = N_z/D = 3$; ou seja, o vetor $u = (4, -2, 3)$.

8.11 SUBMATRIZES, MENORES E MENORES PRINCIPAIS

Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n. Consideremos quaisquer r linhas e r colunas de A. Mais precisamente, consideremos um conjunto $I = (i_1, i_2, \ldots, i_r)$ qualquer de r índices de linha e um conjunto $J = (j_1, j_2, \ldots, j_r)$ qualquer de r índices de coluna. Então I e J definem uma submatriz $r \times r$ de A, denotada por $A(I; J)$, obtida suprimindo todas as linhas e colunas de A cujos índices não estão, respectivamente, em I ou J, respectivamente. Assim,

$$A(I; J) = [a_{st} : s \in I, t \in J]$$

Dizemos que o determinante $|A(I; J)|$ é um *menor* de A de ordem r e

$$(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}|A(I; J)|$$

é o correspondente menor com sinal. (Observe que um menor de ordem $n - 1$ é um menor conforme definimos na Seção 8.7, e o menor com sinal correspondente é um cofator.) Além disso, Se I' e J' denotam, respectivamente, os índices de linhas e colunas restantes, então

$$|A(I'; J')|$$

denota o *menor complementar* e seu sinal (Problema 8.73) é igual ao sinal do menor.

Exemplo 8.13 Seja $A = [a_{ij}]$ uma matriz quadrada de ordem 5 e sejam $I = \{1, 2, 4\}$ e $J = \{2, 3, 5\}$. Então $I' = \{3, 5\}$ e $J' = \{1, 4\}$, e o menor $|M|$ e o menor complementar $|M'|$ são dados a seguir.

$$|M| = |A(I; J)| = \begin{vmatrix} a_{12} & a_{13} & a_{15} \\ a_{22} & a_{23} & a_{25} \\ a_{42} & a_{43} & a_{45} \end{vmatrix} \qquad \text{e} \qquad |M'| = |A(I'; J')| = \begin{vmatrix} a_{31} & a_{34} \\ a_{51} & a_{54} \end{vmatrix}$$

Como $1 + 2 + 4 + 2 + 3 + 5 = 17$ é ímpar, $-|M|$ é o menor com sinal e $-|M'|$ é o menor complementar com sinal.

Menores principais

Dizemos que um menor é *principal* se os índices de linha e coluna são os mesmos ou, equivalentemente, se os elementos diagonais do menor são elementos diagonais da matriz. Observamos que o sinal de um menor principal é sempre $+1$, porque a soma dos índices de linhas e dos mesmos índices de colunas é sempre par.

Exemplo 8.14 Seja $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 5 & 4 \\ -3 & 1 & -2 \end{bmatrix}$. Calcule as somas C_1, C_2 e C_3 dos menores principais de ordens 1, 2 e 3, respectivamente.

(a) Temos três menores principais de ordem 1, a saber,

$$|1| = 1, \qquad |5| = 5, \qquad |-2| = -2, \quad \text{e, portanto,} \quad C_1 = 1 + 5 - 2 = 4$$

Observe que C_1 é, simplesmente, o traço de A, ou seja, $C_1 = \text{tr}(A)$.

(b) Temos três maneiras de escolher dois dentre os três elementos diagonais e cada escolha destas fornece um menor de ordem 2, a saber,

$$\begin{vmatrix} 1 & 2 \\ 3 & 5 \end{vmatrix} = -1, \qquad \begin{vmatrix} 1 & -1 \\ -3 & -2 \end{vmatrix} = 1, \qquad \begin{vmatrix} 5 & 4 \\ 1 & -2 \end{vmatrix} = -14$$

(Observe que esses menores de ordem 2 são os cofatores A_{33}, A_{22} e A_{11} de A, respectivamente.) Assim,

$$C_2 = -1 + 1 - 14 = -14$$

(c) Temos somente uma maneira de escolher três dentre os três elementos diagonais. Logo, o único menor de ordem três é o próprio determinante de A. Assim,

$$C_3 = |A| = -10 - 24 - 3 - 15 - 4 + 12 = -44$$

8.12 MATRIZES EM BLOCOS E DETERMINANTES

O principal resultado desta seção é o teorema seguinte (demonstrado no Problema 8.36).

Teorema 8.12 Seja M uma matriz triangular superior (inferior) em blocos com blocos diagonais $A_1, A_2, \ldots, A_n$. Então

$$\det(M) = \det(A_1)\,\det(A_2)\cdots\det(A_n)$$

Exemplo 8.15 Calcule $|M|$, sendo $M = \left[\begin{array}{cc|ccc} 2 & 3 & 4 & 7 & 8 \\ -1 & 5 & 3 & 2 & 1 \\ \hline 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 3 & -1 & 4 \\ 0 & 0 & 5 & 2 & 6 \end{array}\right]$

Observe que M é uma matriz triangular superior em blocos. Calculamos o determinante de cada bloco diagonal.

$$\begin{vmatrix} 2 & 3 \\ -1 & 5 \end{vmatrix} = 10 + 3 = 13, \qquad \begin{vmatrix} 2 & 1 & 5 \\ 3 & -1 & 4 \\ 5 & 2 & 6 \end{vmatrix} = -12 + 20 + 30 + 25 - 16 - 18 = 29$$

Então $|M| = 13(29) = 377$.

OBSERVAÇÃO Seja $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, em que A, B, C, D são matrizes quadradas. Não é válido em geral que $|M| = |A||D| - |B||C|$. (Ver Problema 8.68.)

8.13 DETERMINANTES E VOLUME

O conceito de determinante está relacionado com os conceitos de área e volume, como segue. Sejam $u_1, u_2, \ldots, u_n$ vetores de $\mathbf{R}^n$. Seja S o paralelepípedo (sólido) determinado por esses vetores, isto é,

$$S = \{a_1u_1 + a_2u_2 + \cdots + a_nu_n : 0 \le a_i \le 1, \text{com } i = 1, \ldots, n\}$$

(No caso $n = 2$, S é um paralelogramo.) Denotemos por $V(S)$ o volume de S (ou a área de S, se $n = 2$). Então

$$V(S) = \text{valor absoluto de } \det(A)$$

onde A é a matriz de linhas $u_1, u_2, \ldots, u_n$. Em geral, $V(S) = 0$ se, e só se, os vetores $u_1, u_2, \ldots, u_n$ não formam um sistema de coordenadas para $\mathbf{R}^n$ (ou seja, se, e só se, os vetores são linearmente dependentes).

Exemplo 8.16 Sejam $u_1 = (1, 1, 0)$, $u_2 = (1, 1, 1)$, $u_3 = (0, 2, 3)$. Encontre o volume $V(S)$ do paralelepípedo S de $\mathbf{R}^3$ (Figura 8-2) determinado por esses três vetores.

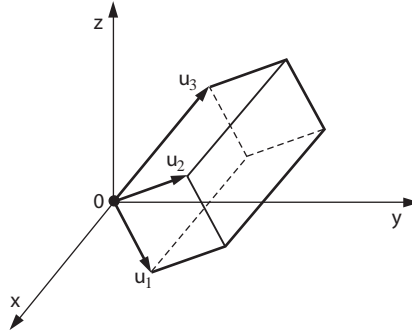



Figura 8-2

Calculamos o determinante da matriz cujas linhas são u_1, u_2, u_3. Temos

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 3 \end{vmatrix} = 3 + 0 + 0 - 0 - 2 - 3 = -2$$

Portanto, $V(S) = |-2| = 2$.

8.14 DETERMINANTE DE UM OPERADOR LINEAR

Seja F um operador linear de um espaço vetorial V de dimensão finita. Seja A a representação matricial de F em relação a alguma base S de V. Então definimos o determinante de F, denotado por det(F), por

$$\det(F) = |A|$$

Se B for uma outra representação matricial de F em relação a alguma outra base S' de V, então A e B são matrizes semelhantes (Teorema 6.7) e $|A| = |B|$ (Teorema 8.7). Em outras palavras, a definição de $\det(F)$ independe da particular escolha da base S de V. (Dizemos, então, que este conceito está *bem definido*).

O teorema seguinte (demonstrado no Problema 8.62) segue de teoremas análogos para matrizes.

Teorema 8.13 Sejam F e G operadores lineares de um espaço vetorial V. Então

(i) $\det(F \circ G) = \det(F)\det(G)$.

(ii) F é invertível se, e só se, $\det(F) \neq 0$.

Exemplo 8.17 Sejam F o operador linear de $\mathbf{R}^3$ dado e A a matriz que representa F na base canônica de $\mathbf{R}^3$.

$$F(x, y, z) = (2x - 4y + z,\ x - 2y + 3z,\ 5x + y - z) \qquad \text{e} \qquad A = \begin{bmatrix} 2 & -4 & 1 \\ 1 & -2 & 3 \\ 5 & 1 & -1 \end{bmatrix}$$

Então

$$\det(F) = |A| = 4 - 60 + 1 + 10 - 6 - 4 = -55$$

8.15 MULTILINEARIDADE E DETERMINANTES

Seja V um espaço vetorial sobre um corpo K e denotemos por $\mathscr{A} = V^n$; o conjunto de todas ênuplas

$$A = (A_1, A_2, ..., A_n)$$

em que os A_i são vetores de V. Temos as definições seguintes.

DEFINIÇÃO Dizemos que uma aplicação $D: \mathscr{A} \to K$ é *multilinear* se for linear em cada componente, como segue.

(i) Se $A_i = B + C$, então

$$D(A) = D(\ldots, B + C, \ldots) = D(\ldots, B, \ldots,) + D(\ldots, C, \ldots)$$

(ii) Se $A_i = kB$, com $k \in K$, então

$$D(A) = D(\ldots, kB, \ldots) = kD(\ldots, B, \ldots)$$

Aplicações multilineares com n componentes também são ditas *n-lineares* e, em particular, com $n = 2$, *bilineares*.

DEFINIÇÃO Dizemos que uma aplicação $D: \mathscr{A} \to K$ é *alternada* se $D(A) = 0$ em cada A com dois elementos iguais, ou seja,

$$D(A_1, A_2, \ldots, A_n) = 0 \qquad \text{se} \qquad A_i = A_j, \quad i \neq j$$

Seja, agora, **M** o conjunto de todas as matrizes quadradas A de ordem n sobre um corpo K. Podemos ver A como uma ênupla consistindo nos vetores linha $A_1, A_2, ..., A_n$ de A, isto é, podemos ver A no formato $A = (A_1, A_2, ..., A_n)$.

O teorema seguinte (demonstrado no Problema 8.37) caracteriza a função determinante.

Teorema 8.14 Existe uma única função D: $\mathbf{M} \to K$ $D: M \to K$ tal que

(i) D é multilinear, (ii) D é alternada, (iii) $D(I) = 1$.

Essa função D é a função determinante, ou seja, $D(A) = |A|$, para qualquer matriz $A \in M$.

Problemas Resolvidos

Cálculo de determinantes

8.1 Calcule o determinante de cada uma das matrizes seguintes.

(a) $A = \begin{bmatrix} 6 & 5 \\ 2 & 3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & -3 \\ 4 & 7 \end{bmatrix}$, (c) $C = \begin{bmatrix} 4 & -5 \\ -1 & -2 \end{bmatrix}$, (d) $D = \begin{bmatrix} t-5 & 6 \\ 3 & t+2 \end{bmatrix}$

Usamos a fórmula $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$:

(a) $|A| = 6(3) - 5(2) = 18 - 10 = 8$
(b) $|B| = 14 + 12 = 26$
(c) $|C| = -8 - 5 = -13$
(d) $|D| = (t-5)(t+2) - 18 = t^2 - 3t - 10 - 18 = t^2 - 10t - 28$

8.2 Calcule o determinante de cada uma das matrizes seguintes.

(a) $A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 4 & 3 \\ 1 & 2 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \\ 1 & 5 & -2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 3 & -5 \\ 3 & -1 & 2 \\ 1 & -2 & 1 \end{bmatrix}$

Usamos o diagrama da Figura 8-1 para obter os seis produtos.

(a) $|A| = 2(4)(1) + 3(3)(1) + 4(2)(5) - 1(4)(4) - 2(3)(2) - 1(3)(5) = 8 + 9 + 40 - 16 - 12 - 15 = 14$

(b) $|B| = -8 + 2 + 30 - 12 + 5 - 8 = 9$

(c) $|C| = -1 + 6 + 30 - 5 + 4 - 9 = 25$

8.3 Calcule o determinante de cada uma das matrizes seguintes.

$$\text{(a)}\quad A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{bmatrix},\ \text{(b)}\quad B = \begin{bmatrix} 4 & -6 & 8 & 9 \\ 0 & -2 & 7 & -3 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix},\ \text{(c)}\quad C = \begin{bmatrix} \frac{1}{2} & -1 & -\frac{1}{3} \\ \frac{3}{4} & \frac{1}{2} & -1 \\ 1 & -4 & 1 \end{bmatrix}.$$

(a) Podemos simplificar as entradas subtraindo da segunda linha o dobro da primeira, isto é, usando a operação sobre as linhas "Substituir R_2 por $-2R_1 + R_2$". Então

$$|A| = \begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 \\ 1 & 0 & -1 \\ 8 & 9 & 1 \end{vmatrix} = 0 - 24 + 36 - 0 + 18 - 3 = 27$$

(b) B é triangular, portanto, $|B|$ = produto das entradas diagonais = –120.

(c) Eliminando as frações, simplificamos as contas. Por isso, multiplicamos a primeira linha R_1 por 6 e a segunda linha R_2 por 4. Então

$$|24C| = \begin{vmatrix} 3 & -6 & -2 \\ 3 & 2 & -4 \\ 1 & -4 & 1 \end{vmatrix} = 6 + 24 + 24 + 4 - 48 + 18 = 28, \quad \text{portanto} \quad |C| = \frac{28}{24} = \frac{7}{6}$$

8.4 Calcule o determinante de cada uma das matrizes seguintes.

$$\text{(a)}\quad A = \begin{bmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{bmatrix},\ \text{(b)}\quad B = \begin{bmatrix} 6 & 2 & 1 & 0 & 5 \\ 2 & 1 & 1 & -2 & 1 \\ 1 & 1 & 2 & -2 & 3 \\ 3 & 0 & 2 & 3 & -1 \\ -1 & -1 & -3 & 4 & 2 \end{bmatrix}$$

(a) Usamos $a_{31} = 1$ como pivô para colocar zeros na primeira coluna, aplicando as operações sobre as linhas "Substituir R_1 por $-2R_3 + R_1$", "Substituir R_2 por $2R_3 + R_2$" e "Substituir R_4 por $R_3 + R_4$". Então

$$|A| = \begin{vmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 1 & -6 \\ 0 & 3 & -2 & -1 \\ 1 & 3 & -2 & 2 \\ 0 & -3 & 2 & 5 \end{vmatrix} = \begin{vmatrix} -1 & 1 & -6 \\ 3 & -2 & -1 \\ -3 & 2 & 5 \end{vmatrix}$$
$$= 10 + 3 - 36 + 36 - 2 - 15 = -4$$

(b) Primeiro reduzimos $|B|$ a um determinante de ordem 4 e depois a um de ordem 3, no qual podemos usar o diagrama da Figura 8-1. Usamos $b_{22} = 1$ como pivô para colocar zeros na segunda coluna, aplicando as operações sobre as linhas "Substituir R_1 por $-2R_2 + R_1$", "Substituir R_3 por $-R_2 + R_3$" e "Substituir R_5 por $R_2 + R_5$". Então

$$|B| = \begin{vmatrix} 2 & 0 & -1 & 4 & 3 \\ 2 & 1 & 1 & -2 & 1 \\ -1 & 0 & 1 & 0 & 2 \\ 3 & 0 & 2 & 3 & -1 \\ 1 & 0 & -2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 4 & 3 \\ -1 & 1 & 0 & 2 \\ 3 & 2 & 3 & -1 \\ 1 & -2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 4 & 5 \\ 0 & 1 & 0 & 0 \\ 5 & 2 & 3 & -5 \\ -1 & -2 & 2 & 7 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 4 & 5 \\ 5 & 3 & -5 \\ -1 & 2 & 7 \end{vmatrix} = 21 + 20 + 50 + 15 + 10 - 140 = -34$$

Cofatores, adjuntas clássicas, inversas

8.5 Seja $A = \begin{bmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{bmatrix}$.

(a) Encontre A_{23}, o cofator (menor com sinal) de 7 em A.

(b) Encontre o menor e o menor com sinal da submatriz $M = A(2, 4; 2, 3)$.

(c) Encontre o menor principal determinado pelas primeira e terceiras entradas diagonais, isto é, por $M = A(1, 3; 1, 3)$.

(a) Tomamos o determinante da submatriz de A obtida suprimindo a linha 2 e a coluna 3 (as que tem o 7) e multiplicamos o determinante por $(-1)^{2+3}$:

$$A_{23} = -\begin{vmatrix} 2 & 1 & 4 \\ 4 & 0 & -3 \\ 3 & -2 & 2 \end{vmatrix} = -(-61) = 61$$

O expoente $2 + 3$ vem dos índices de A_{23}, ou seja, do fato de o 7 aparecer na linha 2 e coluna 3.

(b) Os índices das linhas são 2 e 4 e os das colunas são 2 e 3. Logo, o menor é o determinante

$$|M| = \begin{vmatrix} a_{22} & a_{23} \\ a_{42} & a_{43} \end{vmatrix} = \begin{vmatrix} -4 & 7 \\ -2 & 5 \end{vmatrix} = -20 + 14 = -6$$

e o menor com sinal é $(-1)^{2+4+2+3}|M| = -|M| = -(-6) = 6$.

(c) O menor principal é o determinante

$$|M| = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = \begin{vmatrix} 2 & -3 \\ 4 & 6 \end{vmatrix} = 12 + 12 = 24$$

Observe que as entradas diagonais da submatriz são as entradas diagonais da matriz original. Também, o sinal do menor principal é positivo.

8.6 Seja $B = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 5 & 8 & 9 \end{bmatrix}$. Encontre (a) $|B|$, (b) adj B, (c) B^{-1}, usando adj B.

(a) $|B| = 27 + 20 + 16 - 15 - 32 - 18 = -2$

(b) Queremos a transposta da matriz de cofatores

$$\text{adj } B = \begin{bmatrix} \begin{vmatrix} 3 & 4 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 5 & 9 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 5 & 8 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ 8 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 5 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 5 & 8 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -5 & 2 & 1 \\ -1 & 4 & -3 \\ 1 & -2 & 1 \end{bmatrix}^T = \begin{bmatrix} -5 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -3 & 1 \end{bmatrix}$$

(c) Como $|B| \neq 0$, temos $B^{-1} = \dfrac{1}{|B|}(\text{adj } B) = \dfrac{1}{-2}\begin{bmatrix} -5 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -3 & 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & -2 & 1 \\ -\frac{1}{2} & \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$

8.7 Seja $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 7 & 8 \end{bmatrix}$. Encontre a soma S_k dos menores principais de ordem k de A nos casos

(a) $k = 1$, (b) $k = 2$, (c) $k = 3$.

(a) Os menores principais de ordem 1 são os elementos da diagonal. Assim, S_1 é o traço de A, ou seja,

$$S_1 = \text{tr}(A) = 1 + 5 + 8 = 14$$

(b) Os menores principais de ordem 2 são os cofatores dos elementos diagonais. Assim,

$$S_2 = A_{11} + A_{22} + A_{33} = \begin{vmatrix} 5 & 6 \\ 7 & 8 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 0 & 8 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -2 + 8 - 3 = 3$$

(c) Só há um menor principal de ordem 3, o determinante de A. Então

$$S_3 = |A| = 40 + 0 + 84 - 0 - 42 - 64 = 18$$

8.8 Seja $A = \begin{bmatrix} 1 & 3 & 0 & -1 \\ -4 & 2 & 5 & 1 \\ 1 & 0 & 3 & -2 \\ 3 & -2 & 1 & 4 \end{bmatrix}$. Encontre o número N_k e a soma S_k dos menores principais de ordem

(a) $k = 1$, (b) $k = 2$, (c) $k = 3$, (d) $k = 4$.

Cada subconjunto (não vazio) da diagonal ou, equivalentemente, cada subconjunto (não vazio) de $\{1, 2, 3, 4\}$, determina um menor principal de A, sendo $N_k = \binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ desses de ordem k. Assim,

$$N_1 = \binom{4}{1} = 4, \qquad N_2 = \binom{4}{2} = 6, \qquad N_3 = \binom{4}{3} = 4, \qquad N_4 = \binom{4}{4} = 1$$

(a) $S_1 = |1| + |2| + |3| + |4| = 1 + 2 + 3 + 4 = 10$

(b) $S_2 = \begin{vmatrix} 1 & 3 \\ -4 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 3 & 4 \end{vmatrix} + \begin{vmatrix} 2 & 5 \\ 0 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ -2 & 4 \end{vmatrix} + \begin{vmatrix} 3 & -2 \\ 1 & 4 \end{vmatrix}$

$= 14 + 3 + 7 + 6 + 10 + 14 = 54$

(c) $S_3 = \begin{vmatrix} 1 & 3 & 0 \\ -4 & 2 & 5 \\ 1 & 0 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 3 & -1 \\ -4 & 2 & 1 \\ 3 & -2 & 4 \end{vmatrix} + \begin{vmatrix} 1 & 0 & -1 \\ 1 & 3 & -2 \\ 3 & 1 & 4 \end{vmatrix} + \begin{vmatrix} 2 & 5 & 1 \\ 0 & 3 & -2 \\ -2 & 1 & 4 \end{vmatrix}$

$= 57 + 65 + 22 + 54 = 198$

(d) $S_4 = \det(A) = 378$

Determinantes e sistema de equações lineares

8.9 Use determinantes para resolver o sistema $\begin{cases} 3y + 2x = z + 1 \\ 3x + 2z = 8 - 5y \\ 3z - 1 = x - 2y \end{cases}$.

Primeiro arranjamos a equação no formato padrão, depois calculamos o determinante D da matriz de coeficientes.

$$\begin{aligned} 2x + 3y - z &= 1 \\ 3x + 5y + 2z &= 8 \\ x - 2y - 3z &= -1 \end{aligned} \qquad \text{e} \qquad D = \begin{vmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{vmatrix} = -30 + 6 + 6 + 5 + 8 + 27 = 22$$

Como $D \neq 0$, o sistema tem solução única. Para calcular N_x, N_y, N_z substituímos, respectivamente, os coeficientes de x, y, z na matriz de coeficientes pelos termos constantes. Então

$$N_x = \begin{vmatrix} 1 & 3 & -1 \\ 8 & 5 & 2 \\ -1 & -2 & -1 \end{vmatrix} = 66, \qquad N_y = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 8 & 2 \\ 1 & -1 & -3 \end{vmatrix} = -22, \qquad N_z = \begin{vmatrix} 2 & 3 & 1 \\ 3 & 5 & 8 \\ 1 & -2 & -1 \end{vmatrix} = 44$$

Assim,

$$x = \frac{N_x}{D} = \frac{66}{22} = 3, \qquad y = \frac{N_y}{D} = \frac{-22}{22} = -1, \qquad z = \frac{N_z}{D} = \frac{44}{22} = 2$$

8.10 Considere o sistema $\begin{cases} kx + y + z = 1 \\ x + ky + z = 1 \\ x + y + kz = 1 \end{cases}$

Use determinantes para encontrar os valores de k com os quais o sistema tem
(a) solução única, (b) mais de uma solução, (c) nenhuma solução.

(a) O sistema tem solução única se $D \neq 0$, onde D é o determinante da matriz de coeficientes. Calculamos

$$D = \begin{vmatrix} k & 1 & 1 \\ 1 & k & 1 \\ 1 & 1 & k \end{vmatrix} = k^3 + 1 + 1 - k - k - k = k^3 - 3k + 2 = (k-1)^2(k+2)$$

Assim, o sistema tem solução única se

$$(k-1)^2(k+2) \neq 0, \text{ ou seja, se } k \neq 1 \quad \text{e} \quad k \neq 2$$

(b e c) A eliminação gaussiana mostra que o sistema tem mais de uma solução se $k = 1$ e que o sistema não tem solução se $k = 2$.

Problemas variados

8.11 Encontre o volume $V(S)$ do paralelepípedo S de $\mathbf{R}^3$ determinado pelos vetores dados.

(a) $u_1 = (1, 1, 1), u_2 = (1, 3, -4), u_3 = (1, 2, -5)$.

(b) $u_1 = (1, 2, 4), u_2 = (2, 1, -3), u_3 = (5, 7, 9)$.

$V(S)$ é o valor absoluto do determinante da matriz M cujas linhas são os vetores dados. Assim,

(a) $|M| = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 3 & -4 \\ 1 & 2 & -5 \end{vmatrix} = -15 - 4 + 2 - 3 + 8 + 5 = -7.$ Logo, $V(S) = |-7| = 7$.

(b) $|M| = \begin{vmatrix} 1 & 2 & 4 \\ 2 & 1 & -3 \\ 5 & 7 & 9 \end{vmatrix} = 9 - 30 + 56 - 20 + 21 - 36 = 0.$ Assim, $V(S) = 0$ ou, em outras palavras, u_1, u_2, u_3 são coplanares e linearmente dependentes.

8.12 Encontre $\det(M)$, com $M = \begin{bmatrix} 3 & 4 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 0 & 9 & 2 & 0 & 0 \\ 0 & 5 & 0 & 6 & 7 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} = \left[\begin{array}{cc|c|cc} 3 & 4 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ \hline 0 & 9 & 2 & 0 & 0 \\ \hline 0 & 5 & 0 & 6 & 7 \\ 0 & 0 & 4 & 3 & 4 \end{array}\right]$

M é uma matriz triangular (inferior) em blocos, logo, calculamos o determinante de cada bloco.

$$\begin{vmatrix} 3 & 4 \\ 2 & 5 \end{vmatrix} = 15 - 8 = 7, \qquad |2| = 2, \qquad \begin{vmatrix} 6 & 7 \\ 3 & 4 \end{vmatrix} = 24 - 21 = 3$$

Assim, $|M| = 7(2)(3) = 42$.

8.13 Encontre o determinante do operador $F\colon \mathbf{R}^3 \to \mathbf{R}^3$ dado por

$$F(x, y, z) = (x + 3y - 4z,\ 2y + 7z,\ x + 5y - 3z)$$

O determinante de um operador linear F é igual ao determinante de qualquer matriz que represente F. Por isso, primeiro calculamos a matriz que representa F na base canônica, cujas linhas consistem nos coeficientes de x, y, z, respectivamente. Então

$$A = \begin{bmatrix} 1 & 3 & -4 \\ 0 & 2 & 7 \\ 1 & 5 & -3 \end{bmatrix}, \quad \text{e, portanto,} \quad \det(F) = |A| = -6 + 21 + 0 + 8 - 35 - 0 = -8$$

8.14 Escreva explicitamente $g = g(x_1, x_2, x_3, x_4)$, sendo, em geral,

$$g(x_1, x_2, \ldots, x_n) = \prod_{i<j} (x_i - x_j).$$

O símbolo $\prod$ é usado para um produto de termos, da mesma forma que o símbolo $\sum$ é usado para uma soma de termos. Logo, $\prod_{i<j} (x_i - x_j)$ significa o produto de todos termos $(x_i - x_j)$ para os quais $i < j$. Assim,

$$g = g(x_1, \ldots, x_4) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)$$

8.15 Seja D uma aplicação bilinear alternada. Mostre que $D(A, B) = -D(B, A)$.

Sendo D é alternada, $D(A, A) = 0$, $D(B, B) = 0$. Logo,

$$D(A + B, A + B) = D(A, A) + D(A, B) + D(B, A) + D(B, B) = D(A, B) + D(B, A)$$

No entanto, $D(A + B,\ A + B) = 0$. Assim, $D(A, B) = -D(B, A)$, como queríamos mostrar.

Permutações

8.16 Encontre a paridade (sinal) da permutação $\sigma = 364152$.

Contamos o número de inversões, ou seja, para cada elemento k, contamos o número de elementos i de σ tais que $i > k$ e i precede k em σ. Temos

$k = 1$: 3 números (3, 6, 4) $\qquad k = 4$: 1 número (6)
$k = 2$: 4 números (3, 6, 4, 5) $\qquad k = 5$: 1 número (6)
$k = 3$: 0 números $\qquad k = 6$: 0 números

Como $3 + 4 + 0 + 1 + 1 + 0 = 9$ é ímpar, σ é uma permutação ímpar e $\sigma = -1$.

8.17 Sejam $\sigma = 24513$ e $\tau = 41352$ permutações de S_5. Encontre (a) $\tau \circ \sigma$, (b) σ^{-1}.

Lembre que $\sigma = 24513$ e $\tau = 41352$ são maneiras concisas de escrever

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix} \quad \text{ou} \quad \sigma(1) = 2, \quad \sigma(2) = 4, \quad \sigma(3) = 5, \quad \sigma(4) = 1, \quad \sigma(5) = 3$$

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2ç \end{pmatrix} \quad \text{ou} \quad \tau(1) = 4, \quad \tau(2) = 1, \quad \tau(3) = 3, \quad \tau(4) = 5, \quad \tau(5) = 2$$

(a) Efetuando σ e depois τ em 1, 2, 3, 4, 5, obtemos

$$1 \to 2 \to 1, \qquad 2 \to 4 \to 5, \qquad 3 \to 5 \to 2, \qquad 4 \to 1 \to 4, \qquad 5 \to 3 \to 3$$

[Isto é, por exemplo, $(\tau \circ \sigma)(1) = \tau(\sigma(1)) = \tau(2) = 1$.] Assim, $\tau \circ \sigma = 15243$.

(b) Por definição, $\sigma^{-1}(j) = k$ se, e só se, $\sigma(k) = j$. Logo,

$$\sigma^{-1} = \begin{pmatrix} 2 & 4 & 5 & 1 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix} \quad \text{ou} \quad \sigma^{-1} = 41523$$

8.18 Seja $\sigma = j_1 j_2 \ldots j_n$ uma permutação qualquer de S_n. Mostre que, para cada inversão (i, k) com $i > k$, mas i precedendo k em σ, existe um par (i^*, j^*) tal que

$$i^* < k^* \qquad \text{e} \qquad \sigma(i^*) > \sigma(j^*) \tag{1}$$

e vice-versa. Assim, σ é par ou ímpar, dependendo de existir um número par ou ímpar de pares satisfazendo (1).

Escolhemos i^* e k^* dados por $\sigma(i^*) = i$ e $\sigma(k^*) = k$. Então $i > k$ se, e só se, $\sigma(i^*) > \sigma(k^*)$, e i precede k em σ se, e só se, $i^* < k^*$.

8.19 Considere os polinômios $g = g(x_1, \dots, x_n)$ e $\sigma(g)$, definidos por

$$g = g(x_1, \dots, x_n) = \prod_{i<j}(x_i - x_j) \qquad \text{e} \qquad \sigma(g) = \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)})$$

(Ver Problema 8.14.) Mostre que $\sigma(g) = g$ se σ for uma permutação par e $\sigma(g) = -g$, se σ for uma permutação ímpar, ou seja, $\sigma(g) = (\text{sgn } \sigma)g$.

Como σ é injetora e sobre,

$$\sigma(g) = \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = \prod_{i<j \text{ ou } i>j}(x_i - x_j)$$

Assim, $\sigma(g)$ ou $\sigma(g) = -g$, dependendo de existir um número par ou ímpar de termos da forma $x_i - x_j$, com $i > j$. Observe que, para cada par (i, j) para o qual

$$i < j \qquad \text{e} \qquad \sigma(i) > \sigma(j) \tag{1}$$

existe um termo $(x_{\sigma(i)} - x_{\sigma(j)})$ em $\sigma(g)$ para o qual $\sigma(i) > \sigma(j)$. Como σ é par se, e só se, existir um número par de pares satisfazendo (1), temos $\sigma(g) = g$ se, e só se, σ é par. Assim, $\sigma(g) = -g$ se, e só se, σ é ímpar.

8.20 Sejam $\sigma, \tau \in S_n$. Mostre que $\text{sgn}(\tau \circ \sigma) = (\text{sgn } \tau)(\text{sgn } \sigma)$. Assim, o produto de duas permutações pares ou ímpares é par e o produto de uma permutação par por uma ímpar é ímpar.

Usando o Problema 8.19, temos

$$\text{sgn}(\tau \circ \sigma)\, g = (\tau \circ \sigma)(g) = \tau(\sigma(g)) = \tau((\text{sgn } \sigma)g) = (\text{sgn } \tau)(\text{sgn } \sigma)g$$

Por isso, sgn $(\tau \circ \sigma) = (\text{sgn } \tau)(\text{sgn } \sigma)$.

8.21 Considere a permutação $\sigma = j_1 j_2 \cdots j_n$. Mostre que $\sigma^{-1} = \text{sgn } \sigma$ e que, dados escalares a_{ij}, mostre que

$$a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} = a_{1k_1} a_{2k_2} \cdots a_{nk_n}$$

onde $\sigma^{-1} = k_1 k_2 \cdots k_n$.

Temos $\sigma^{-1} \circ \sigma = \varepsilon$, a permutação identidade. Como ε é par, σ^{-1} e σ são, ambas, pares ou ímpares. Assim, sgn $\sigma^{-1} = \text{sgn } \sigma$.

Como $\sigma = j_1 j_2 \cdots j_n$ é uma permutação, $a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} = a_{1k_1} a_{2k_2} \cdots a_{nk_n}$. Então $k_1, k_2, \dots, k_n$ tem a propriedade

$$\sigma(k_1) = 1, \qquad \sigma(k_2) = 2, \qquad \dots, \qquad \sigma(k_n) = n$$

Seja $\tau = k_1 k_2 \cdots k_n$. Então, para $i = 1, \dots, n$,

$$(\sigma \circ \tau)(i) = \sigma(\tau(i)) = \sigma(k_i) = i$$

Logo, $\sigma \circ \tau = \varepsilon$, a permutação identidade. Assim, $\tau = \sigma^{-1}$.

Demonstrações dos teoremas

8.22 Demonstre o Teorema 8.1. $|A^T| = |A|$.

Se $A = [a_{ij}]$, então $A^T = [b_{ij}]$, com $b_{ij} = a_{ij}$. Logo,

$$|A^T| = \sum_{\sigma \in S_n} (\text{sgn } \sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} (\text{sgn } \sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

Seja $\tau = \sigma^{-1}$. Pelo Problema 8.21, sgn $\tau = \text{sgn } \sigma$, e $a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} = a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$. Logo,

$$|A^T| = \sum_{\sigma \in S_n} (\text{sgn } \tau) a_{1t(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$$

Contudo, quando σ percorre todos os elementos de S_n, $\tau = \sigma^{-1}$ também percorre todos os elementos de S_n. Assim, $|A^T| = |A|$.

8.23 Demonstre o Teorema 8.3(i). Se duas linhas (colunas) de A trocaram de posição, então $|B| = -|A|$.

Provamos a afirmação no caso em que foram trocadas duas colunas. Seja τ a transposição que troca os dois índices correspondente às duas colunas de A que foram trocadas. Se $A = [a_{ij}]$ e $B = [b_{ij}]$, então $b_{ij} = a_{i\tau(j)}$. Logo, dada qualquer permutação σ,

$$b_{1\sigma(1)}b_{2\sigma(2)}\cdots b_{n\sigma(n)} = a_{1(\tau\circ\sigma)(1)}a_{2(\tau\circ\sigma)(2)}\cdots a_{n(\tau\circ\sigma)(n)}$$

Assim,

$$|B| = \sum_{\sigma\in S_n}(\operatorname{sgn}\sigma)b_{1\sigma(1)}b_{2\sigma(2)}\cdots b_{n\sigma(n)} = \sum_{\sigma\in S_n}(\operatorname{sgn}\sigma)a_{1(\tau\circ\sigma)(1)}a_{2(\tau\circ\sigma)(2)}\cdots a_{n(\tau\circ\sigma)(n)}$$

Como a transposição τ é uma permutação ímpar, $\operatorname{sgn}(\tau\circ\sigma) = (\operatorname{sgn}\tau)(\operatorname{sgn}\sigma) = -\operatorname{sgn}\sigma$. Por isso, $\operatorname{sgn}\sigma = -\operatorname{sgn}(\tau\circ\sigma)$, e, portanto,

$$|B| = -\sum_{\sigma\in S_n}[\operatorname{sgn}(\tau\circ\sigma)]a_{1(\tau\circ\sigma)(1)}a_{2(\tau\circ\sigma)(2)}\cdots a_{n(\tau\circ\sigma)(n)}$$

Contudo, quando σ percorre todos os elementos de S_n, $\tau\circ\sigma$ também percorre todos os elementos de S_n. Assim, $|B| = -|A|$.

8.24 Demonstre o Teorema 8.2. Seja A uma matriz quadrada.

(i) Se A tem uma linha (coluna) de zeros, então $|A| = 0$.

(i) Se A tem duas linhas (colunas) idênticas, então $|A| = 0$.

(iii) Se A é triangular, então $|A|$ = produto dos elementos diagonais. Assim, $|I| = 1$.

(i) Cada parcela de $|A|$ contém um fator de cada linha e, portanto, da linha de zeros. Assim, cada parcela de $|A|$ é nula e, assim, $|A| = 0$.

(ii) Suponha que $1 + 1 \neq 0$ em K. Trocando as duas linhas idênticas de A entre si, continuamos com a mesma matriz A. Logo, pelo Problema 8.23, $|A| = -|A|$ e, portanto, $|A| = 0$.

Agora, suponha que $1 + 1 = 0$ em K. Então $\operatorname{sgn}\sigma = 1$, para cada $\sigma \in S_n$. Como A tem duas linhas idênticas, podemos rearranjar as parcelas de |A| em pares de parcelas iguais. Como cada par é 0, o determinante de A é 0.

(iii) Suponha $A = [a_{ij}]$ seja triangular inferior, ou seja, as entradas acima da diagonal são todas nulas, $a_{ij} = 0$ se $i < j$. Considere uma parcela t do determinante de A.

$$t = (\operatorname{sgn}\sigma)a_{1i_1}a_{2i_2}\cdots a_{ni_n}, \qquad \text{onde} \qquad \sigma = i_1 i_2 \cdots i_n$$

Suponha que $i_1 \neq 1$. Então $1 < i_1$ e, portanto, $a_{1i_1} = 0$. Logo, $t = 0$. Ou seja, é nulo cada termo em que $i_1 \neq 1$.

Agora suponha que $i_1 = 1$, mas $i_2 \neq 2$. Então $2 < i_2$, e, portanto, $a_{2i_2} = 0$. Logo, $t = 0$. Ou seja, é nulo cada termo em que $i_1 \neq 1$ ou $i_2 \neq 2$.

Analogamente, é nulo cada termo em que $i_1 \neq 1$ ou $i_2 \neq 2$ ou ... ou $i_n \neq n$. Por isso, $|A| = a_{11}a_{22}\cdots a_{nn}$ = produto dos elementos diagonais.

8.25 Demonstre o Teorema 8.3. B é obtida a partir de A por uma operação elementar.

(i) Se duas linhas (colunas) trocaram de posição, então $|B| = -|A|$.

(ii) Se uma linha (coluna) de A foi multiplicada por um escalar k, então $|B| = k|A|$.

(iii) Se um múltiplo de uma linha (coluna) de A foi somado a uma outra linha (coluna) de A, então $|B| = |A|$.

(i) Esse resultado está provado no Problema 8.23.

(ii) Se a j-ésima linha de A for multiplicada por k, então cada parcela de $|A|$ será multiplicada por k, portanto, $|B| = k|A|$. Ou seja,

$$|B| = \sum_{\sigma}(\operatorname{sgn}\sigma)a_{1i_1}a_{2i_2}\cdots(ka_{ji_j})\cdots a_{ni_n} = k\sum_{\sigma}(\operatorname{sgn}\sigma)a_{1i_1}a_{2i_2}\cdots a_{ni_n} = k|A|$$

(iii) Suponha que à j-ésima linha de A somamos c vezes a k-ésima linha de A. Usando o símbolo chapéu ^ para denotar a j-ésima posição numa parcela do determinante, temos

$$\begin{aligned}|B| &= \sum_{\sigma} (\text{sgn } \sigma) a_{1i_1} a_{2i_2} \cdots (\widehat{ca_{ki_k} + a_{ji_j}}) \ldots a_{ni_n} \\ &= c \sum_{\sigma} (\text{sgn } \sigma) a_{1i_1} a_{2i_2} \cdots \widehat{a_{ki_k}} \cdots a_{ni_n} + \sum_{\sigma} (\text{sgn } \sigma) a_{1i_1} a_{2i_2} \cdots a_{ji_j} \cdots a_{ni_n}\end{aligned}$$

A primeira soma é o determinante de uma matriz cujas k-ésima e j-ésima linhas são iguais. Assim, pelo Teorema 8.2(ii), essa soma é nula. A segunda soma é o determinante de A. Logo, $|B| = c \cdot 0 + |A| = |A|$.

8.26 Demonstre o Lema 8.6. Seja E uma matriz elementar. Então, $|EA| = |E||A|$.

Considere as operações elementares com as linhas: (i) multiplicar uma linha por uma constante $k \neq 0$, (ii) trocar duas linhas entre si, (iii) somar um múltiplo de uma linha a uma outra linha.

Sejam E_1, E_2, E_3 as matrizes elementares correspondentes. Ou seja, E_1, E_2, E_3 são obtidas aplicando as operações dadas à matriz identidade I. Pelo Problema 8.25,

$$|E_1| = k|I| = k, \qquad |E_2| = -|I| = -1, \qquad |E_3| = |I| = 1$$

Sabemos (Teorema 3.11) que E_iA é a matriz obtida aplicando a operação correspondente a A. Assim, pelo Teorema 8.3, obtemos as afirmações do lema, como segue.

$$|E_1A| = k|A| = |E_1||A|, \qquad |E_2A| = -|A| = |E_2||A|, \qquad |E_3A| = |A| = 1|A| = |E_3||A|$$

8.27 Seja B uma matriz equivalente por linhas a uma matriz quadrada A. Prove que $|B| = 0$ se, e só se, $|A| = 0$.

Pelo Teorema 8.3, o efeito de uma operação elementar é trocar o sinal do determinante ou multiplicar o determinante por um escalar não nulo. Assim, $|B| = 0$ se, e só se, $|A| = 0$.

8.28 Demonstre o Teorema 8.5. Seja A uma matriz quadrada. São equivalentes as afirmações dadas.

(i) A é invertível, (ii) $AX = 0$ possui apenas a solução nula, (iii) $\det(A) \neq 0$.

A demonstração utiliza eliminação gaussiana. Se A for invertível, então A é equivalente por linhas a I. Como $|I| \neq 0$, o Problema 8.27 garante que $|A| \neq 0$. Se A não é invertível, então A é equivalente por linhas a uma matriz com uma linha nula. Logo, $\det(A) = 0$. Assim, (i) é equivalente a (iii).
Se $AX = 0$ tem somente a solução nula $X = 0$, então A é equivalente por linhas a I e A é invertível. Reciprocamente, se A é invertível, com inversa A^{-1}, então

$$X = IX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}0 = 0$$

é a única solução de $AX = 0$. Assim, (i) é equivalente a (ii).

8.29 Demonstre o Teorema 8.4. $|AB| = |A||B|$.

Se A é não singular, então AB também é não singular e, portanto, $|AB| = 0 = |A||B|$. Por outro lado, se A é não singular, então $A = E_n \cdots E_2E_1$, é um produto de matrizes elementares. Logo, pelo Lema 8.6 e indução, obtemos

$$|AB| = |E_n \cdots E_2E_1B| = |E_n| \cdots |E_2||E_1||B| = |A||B|$$

8.30 Seja P invertível. Prove que $|P^{-1}| = |P|^{-1}$.

$P^{-1}P = I$. Logo, $1 = |I| = |P^{-1}P| = |P^{-1}||P|$, e, portanto, $|P^{-1}| = |P|^{-1}$.

8.31 Demonstre o Teorema 8.7. Se A e B são matrizes semelhantes, então $|A| = |B|$.

Se A e B são semelhantes, existe uma matriz invertível P tal que $B = P^{-1}AP$. Logo, usando o Problema 8.30, obtemos $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A||P^{-1}||P = |A|$.
Observamos que, embora as matrizes P^{-1} e A possam não comutar, seus determinantes $|P^{-1}|$ e $|A|$ comutam, por serem escalares do corpo K.

8.32 Demonstre o Teorema 8.8 (Laplace). Sejam $A = [a_{ij}]$ e A_{ij} o cofator de a_{ij}. Então, para quaisquer i ou j,

$$|A| = a_{i1}A_{i1} + \cdots + a_{in}A_{in} \qquad \text{e} \qquad |A| = a_{1j}A_{1j} + \cdots + a_{nj}A_{nj}$$

Como $|A| = |A^T|$, basta provar uma dessas expansões, digamos, a primeira, em termos das linhas de A. Cada parcela de $|A|$ contém uma, e só uma, entrada da i-ésima linha $(a_{i1}, a_{i2}, \ldots, a_{in})$ de A. Logo, podemos escrever $|A|$ na forma

$$|A| = a_{i1}A^*_{i1} + a_{i2}A^*_{i2} + \cdots + a_{in}A^*_{in}$$

(Observe que A^*_{ij} é uma soma de parcelas que não tem elemento algum da i-ésima linha de A.) Assim, o teorema estará provado se conseguirmos mostrar que

$$A^*_{ij} = A_{ij} = (-1)^{i+j}|M_{ij}|$$

onde M_{ij} é a matriz obtida suprimindo a linha e a coluna que contêm a entrada a_{ij}. (Historicamente, a expressão A^*_{ij} era denominada cofator de a_{ij} e, portanto, o teorema reduz a mostrar que as duas definições de cofator são equivalentes.)

Primeiro consideramos o caso em que $i = n$, $j = n$. Então a soma das parcelas de $|A|$ contendo a_{nn} é

$$a_{nn}A^*_{nn} = a_{nn}\sum_{\sigma}(\text{sgn }\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n-1,\sigma(n-1)}$$

onde somamos sobre todas as permutações $\sigma \in S_n$ para as quais $\sigma(n) = n$. Contudo, isso é equivalente (mostre isso!) a somar sobre todas as permutações de $\{1, ..., n-1\}$. Assim, $A^*_{nn} = |M_{nn}| = (-1)^{n+n}|M_{nn}|$.
Agora, consideramos i e j arbitrários. Trocamos a i-ésima linha com cada linha que a sucede até ela ser a última e, em seguida, trocamos a j-ésima coluna com cada coluna que a sucede até ela ser a última. Observe que o determinante $|M_{ij}|$ não se altera, porque as posições relativas das outras linhas e colunas não são afetadas por essas trocas. Contudo, o "sinal" de $|A|$ e de A^*_{ij} foi trocado $n-i$ vezes e, depois, outras $n-j$ vezes. Em vista disso,

$$A^*_{ij} = (-1)^{n-i+n-j}|M_{ij}| = (-1)^{i+j}|M_{ij}|$$

8.33 Seja $A = [a_{ij}]$ e seja B a matriz obtida de A substituindo a i-ésima linha de A pelo vetor coluna $(b_{i1}, \ldots, b_{in})$ Mostre que

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

Além disso, mostre que, para $j \neq i$,

$$a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0 \qquad \text{e} \qquad a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$$

Seja $B = [b_{ij}]$. Pelo Teorema 8.8,

$$|B| = b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in}$$

Como B_{ij} não depende da i-ésima linha de B, obtemos $B_{ij} = A_{ij}$ para $j = 1, ..., n$. Logo,

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

Seja, agora, A' a matriz obtida a partir de A substituindo a i-ésima linha de A pela j-ésima linha de A. Como A' tem duas linhas iguais, $|A'| = 0$. Assim, pelo resultado que acabamos de ver, obtemos

$$|A'| = a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0$$

Usando que $|A| = |A^T|$, também obtemos $a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$.

8.34 Demonstre o Teorema 8.9. $A(\text{adj } A) = (\text{adj } A)A = |A|I$.

Sejam $A = [a_{ij}]$ e $A(\text{adj } A) = [b_{ij}]$. A i-ésima linha de A é

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \tag{1}$$

Como adj A é a transposta da matriz de cofatores, a j-ésima coluna de adj A é a transposta dos cofatores da j-ésima linha de A, como segue.

$$(A_j, A_{j2}, \ldots, A_{jn})^T \tag{2}$$

Agora, a ij-ésima entrada b_{ij} de $A(\text{adj } A)$ é obtida multiplicando as expressões (1) e (2), como segue.

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

Pelo Teorema 8.8 e o Problema 8.33,

$$b_{ij} = \begin{cases} |A| & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$$

Por isso, $A(\text{adj } A)$ é a matriz diagonal com elementos diagonais iguais a $|A|$. Em outras palavras, $A(\text{adj } A) = |A|I$. Analogamente, $(\text{adj } A)A = |A|I$.

8.35 Demonstre o Teorema 8.10 (Regra de Cramer). O sistema (quadrado) $AX = B$ possui uma única solução se, e só se, $D \neq 0$. Nesse caso, $x_i = N_i/D$, para cada i.

Já vimos que $AX = B$ possui uma única solução se, e só se, A é invertível e que A é invertível se, e só se, $D = |A| \neq 0$. Suponha, agora, que $D \neq 0$. Pelo Teorema 8.9, $A^{-1} = (1/D)(\text{adj } A)$. multiplicando $AX = B$ por A^{-1} obtemos

$$X = A^{-1}AX = (1/D)(\text{adj } A)B \tag{1}$$

Observe que a i-ésima linha de $(1/D)(\text{adj } A)$ é $(1/D)(A_{1i}, A_{2i}, \ldots, A_{ni})$. Se $B = (b_1, b_2, \ldots, b_n)^T$, então, por (1), temos

$$x_i = (1/D)(b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni})$$

No entanto, como no Problema 8.33, temos $b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni} = N_i$, que é o determinante da matriz obtida substituindo a i-ésima coluna de A pelo vetor coluna B. Assim, $x_i = (1/D)N_i$, como queríamos mostrar.

8.36 Demonstre o Teorema 8.12. Seja M uma matriz triangular superior (inferior) em blocos com blocos diagonais $A_1, A_2, \ldots, A_n$. Então

$$\det(M) = \det(A_1)\,\det(A_2)\cdots\det(A_n)$$

Basta provar o teorema para $n = 2$, ou seja, quando M é uma matriz quadrada da forma $M = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$. A demonstração no caso geral segue facilmente, por indução.

Sejam $A = [a_{ij}]$ uma matriz quadrada de ordem r, $B = [b_{ij}]$ uma matriz quadrada de ordem s e $M = [m_{ij}]$ uma matriz quadrada de ordem n, com $n = r + s$. Por definição,

$$\det(M) = \sum_{\sigma \in S_n} (\text{sgn } \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}$$

Se $i > r$ e $j \leq r$, então $m_{ij} = 0$. Assim, basta considerar aquelas permutações σ tais que

$$\sigma\{r+1, r+2, \ldots, r+s\} = \{r+1, r+2, \ldots, r+s\} \qquad \text{e} \qquad \sigma\{1, 2, \ldots, r\} = \{1, 2, \ldots, r\}$$

Seja $\sigma_1(k) = \sigma(k)$ para $k \leq r$, e seja $\sigma_2(k) = \sigma(r+k) - r$ para $k \leq s$. Então

$$(\text{sgn } \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)} = (\text{sgn } \sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} \cdots a_{r\sigma_1(r)} (\text{sgn } \sigma_2) b_{1\sigma_2(1)} b_{2\sigma_2(2)} \cdots b_{s\sigma_2(s)}$$

que implica $\det(M) = \det(A)\,\det(B)$.

8.37 Demonstre o Teorema 8.14. Existe uma única função $D : \mathbf{M} \to K$ tal que

(i) D é multilinear, (ii) D é alternada, (iii) $D(I) = 1$.

Essa função D é a função determinante, ou seja, $D(A) = |A|$.

Seja D a função determinante $D(A) = |A|$. Devemos mostrar que D satisfaz (i), (ii) e (iii) e que D é única função que satisfaz (i), (ii) e (iii).

Pelo Teorema 8.2, D satisfaz (ii) e (iii). Mostremos que é multilinear. Suponha que a i-ésima linha de $A = [a_{ij}]$ seja da forma $(b_{i1} + c_{i1},\ b_{i2} + c_{i2},\ \ldots,\ b_{in} + c_{in})$.. Então

$$\begin{aligned} D(A) &= D(A_1,\ \ldots,\ B_i + C_i,\ \ldots,\ A_n) \\ &= \sum_{S_n} (\text{sgn } \sigma) a_{1\sigma(1)} \cdots a_{i-1,\sigma(i-1)} (b_{i\sigma(i)} + c_{i\sigma(i)}) \cdots a_{n\sigma(n)} \\ &= \sum_{S_n} (\text{sgn } \sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{S_n} (\text{sgn } \sigma) a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)} \\ &= D(A_1, \ldots, B_i, \ldots, A_n) + D(A_1, \ldots, C_i, \ldots, A_n) \end{aligned}$$

Também, pelo Teorema 8.3(ii),

$$D(A_1, \ldots, kA_i, \ldots, A_n) = kD(A_1, \ldots, A_i, \ldots, A_n)$$

Assim, D é multilinear, ou seja, D satisfaz (i).
Demonstremos, agora, a unicidade de D. Seja D uma aplicação que satisfaz (i), (ii) e (iii). Se $\{e_1, ..., e_n\}$ é a base canônica de K^n, então, por (iii), $D(e_1, e_2, \ldots, e_n) = D(I) = 1$. Usando (ii), também temos

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = \text{sgn}\ \sigma, \qquad \text{onde} \qquad \sigma = i_1 i_2 \cdots i_n \tag{1}$$

Seja, agora, $A = [a_{ij}]$. Observe que a k-ésima linha A_k de A é

$$A_k = (a_{k1}, a_{k2}, \ldots, a_{kn}) = a_{k1}e_1 + a_{k2}e_2 + \cdots + a_{kn}e_n$$

Assim,

$$D(A) = D(a_{11}e_1 + \cdots + a_{1n}e_n,\ a_{21}e_1 + \cdots + a_{2n}e_n,\ \ldots,\ a_{n1}e_1 + \cdots + a_{nn}e_n)$$

Usando a multilinearidade de D, podemos escrever $D(A)$ como uma soma de parcelas do tipo

$$\begin{aligned} D(A) &= \sum D(a_{1i_1}e_{i_1}, a_{2i_2}e_{i_2}, \ldots, a_{ni_n}e_{i_n}) \\ &= \sum (a_{1i_1}a_{2i_2} \cdots a_{ni_n}) D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \end{aligned} \tag{2}$$

onde a soma percorre todas as sequências $i_1 i_2 \ldots i_n$, com $i_k \in \{1, \ldots, n\}$. Se dois dos índices forem iguais, digamos, $i_j = i_k$ com $j \neq k$, então, por (ii),

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = 0$$

Dessa forma, a soma de (2) só precisa ser somada sobre todas as permutações $\sigma = i_1 i_2 \cdots i_n$. Usando (1), finalmente obtemos que

$$\begin{aligned} D(A) &= \sum_\sigma (a_{1i_1}a_{2i_2} \cdots a_{ni_n}) D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \\ &= \sum_\sigma (\text{sgn}\ \sigma) a_{1i_1}a_{2i_2} \cdots a_{ni_n}, \qquad \text{onde} \qquad \sigma = i_1 i_2 \cdots i_n \end{aligned}$$

Assim, D é a função determinante e o teorema está provado.

Problemas Complementares

Cálculo de determinantes

8.38 Calcule

(a) $\begin{vmatrix} 2 & 6 \\ 4 & 1 \end{vmatrix}$, (b) $\begin{vmatrix} 5 & 1 \\ 3 & -2 \end{vmatrix}$, (c) $\begin{vmatrix} -2 & 8 \\ -5 & -3 \end{vmatrix}$, (d) $\begin{vmatrix} 4 & 9 \\ 1 & -3 \end{vmatrix}$, (e) $\begin{vmatrix} a+b & a \\ b & a+b \end{vmatrix}$

8.39 Encontre todos os t tais que (a) $\begin{vmatrix} t-4 & 3 \\ 2 & t-9 \end{vmatrix} = 0$, (b) $\begin{vmatrix} t-1 & 4 \\ 3 & t-2 \end{vmatrix} = 0$

8.40 Calcule o determinante de cada uma das matrizes seguintes.

(a) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{bmatrix}$, (b) $\begin{bmatrix} 3 & -2 & -4 \\ 2 & 5 & -1 \\ 0 & 6 & 1 \end{bmatrix}$, (c) $\begin{bmatrix} -2 & -1 & 4 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{bmatrix}$, (d) $\begin{bmatrix} 7 & 6 & 5 \\ 1 & 2 & 1 \\ 3 & -2 & 1 \end{bmatrix}$

8.41 Encontre o determinante de cada uma das matrizes seguintes.

$$\text{(a)}\ \begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 3 & -1 & 1 & -2 \\ 4 & -3 & 0 & 2 \end{bmatrix}, \quad \text{(b)}\ \begin{bmatrix} 2 & 1 & 3 & 2 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{bmatrix}$$

8.42 Calcule

$$\text{(a)}\ \begin{vmatrix} 2 & -1 & 3 & -4 \\ 2 & 1 & -2 & 1 \\ 3 & 3 & -5 & 4 \\ 5 & 2 & -1 & 4 \end{vmatrix}, \quad \text{(b)}\ \begin{vmatrix} 2 & -1 & 4 & -3 \\ -1 & 1 & 0 & 2 \\ 3 & 2 & 3 & -1 \\ 1 & -2 & 2 & -3 \end{vmatrix}, \quad \text{(c)}\ \begin{vmatrix} 1 & -2 & 3 & -1 \\ 1 & 1 & -2 & 0 \\ 2 & 0 & 4 & -5 \\ 1 & 4 & 4 & -6 \end{vmatrix}$$

8.43 Calcule cada um dos determinantes seguintes.

$$\text{(a)}\ \begin{vmatrix} 1 & 2 & -1 & 3 & 1 \\ 2 & -1 & 1 & -2 & 3 \\ 3 & 1 & 0 & 2 & -1 \\ 5 & 1 & 2 & -3 & 4 \\ -2 & 3 & -1 & 1 & -2 \end{vmatrix}, \quad \text{(b)}\ \begin{vmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 6 & 2 \\ 0 & 0 & 2 & 3 & 1 \end{vmatrix}, \quad \text{(c)}\ \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 6 & 5 & 1 \\ 0 & 0 & 0 & 7 & 4 \\ 0 & 0 & 0 & 2 & 3 \end{vmatrix}$$

Cofatores, adjuntas clássicas, inversas

8.44 Encontre $\det(A)$, adj A e A^{-1}, com

$$\text{(a)}\quad A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix}, \quad \text{(b)}\quad A = \begin{bmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

8.45 Encontre a adjunta clássica de cada matriz do Problema 8.41.

8.46 Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. (a) Encontre adj A, (b) Mostre que adj(adj A) $= A$, (c) Quando temos $A = \text{adj}\ A$?

8.47 Mostre que, se A é diagonal (triangular), então adj A é diagonal (triangular).

8.48 Seja $A = [a_{ij}]$ uma matriz triangular. Mostre que

(a) A é invertível se, e só se, cada entrada diagonal $a_{ii} \neq 0$.

(b) Os elementos diagonais de A^{-1} (caso exista) são os recíprocos a_{ii}^{-1}, dos elementos diagonais de A.

Menores, menores principais

8.49 Sejam $A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 0 & -2 & 3 \\ 3 & -1 & 2 & 5 \\ 4 & -3 & 0 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 2 & -3 & 1 & 4 \\ 0 & -5 & 2 & 1 \\ 3 & 0 & 5 & -2 \end{bmatrix}$. Encontre o menor e o menor com sinal correspondente às submatrizes dadas.

(a) $A(1, 4;\ \ 3, 4)$, (b) $B(1, 4;\ \ 3, 4)$, (c) $A(2, 3;\ \ 2, 4)$, (d) $B(2, 3;\ \ 2, 4)$.

8.50 Para $k = 1, 2, 3$, encontre a soma S_k de todos os menores principais de ordem k, com

$$\text{(a)}\quad A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & -4 & 3 \\ 5 & -2 & 1 \end{bmatrix}, \quad \text{(b)}\quad B = \begin{bmatrix} 1 & 5 & -4 \\ 2 & 6 & 1 \\ 3 & -2 & 0 \end{bmatrix}, \quad \text{(c)}\quad C = \begin{bmatrix} 1 & -4 & 3 \\ 2 & 1 & 5 \\ 4 & -7 & 11 \end{bmatrix}$$

8.51 Para $k = 1, 2, 3, 4$, encontre a soma S_k de todos os menores principais de ordem k, com

$$\text{(a)}\quad A = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 1 & -2 & 0 & 5 \\ 0 & 1 & -2 & 2 \\ 4 & 0 & -1 & -3 \end{bmatrix}, \quad \text{(b)}\quad B = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 0 & 4 \\ 2 & 7 & 4 & 5 \end{bmatrix}$$

Determinantes e sistemas de equações lineares

8.52 Resolva os sistemas dados com determinantes.

$$\text{(a)}\quad \begin{cases} 3x + 5y = 8 \\ 4x - 2y = 1 \end{cases}, \quad \text{(b)}\quad \begin{cases} 2x - 3y = -1 \\ 4x + 7y = -1 \end{cases}, \quad \text{(c)}\quad \begin{cases} ax - 2by = c \\ 3ax - 5by = 2c \end{cases} \quad (ab \neq 0)$$

8.53 Resolva os sistemas dados com determinantes.

$$\text{(a)}\quad \begin{cases} 2x - 5y + 2z = 2 \\ x + 2y - 4z = 5 \\ 3x - 4y - 6z = 1 \end{cases}, \quad \text{(b)}\quad \begin{cases} 2z + 3 = y + 3x \\ x - 3z = 2y + 1 \\ 3y + z = 2 - 2x \end{cases}$$

8.54 Demonstre o Teorema 8.11. Um sistema $AX = 0$ tem alguma solução não nula se, e só se, $D = |A| = 0$.

Permutações

8.55 Encontre as paridades das permutações $\sigma = 32154$, $\tau = 13524$, $\pi = 42531$ de S_5.

8.56 Para as permutações do Problema 8.55, encontre

(a) $\tau \circ \sigma$, (b) $\pi \circ \sigma$, (c) σ^{-1}, (d) τ^{-1}.

8.57 Seja $\tau \in S_n$. Mostre que $\tau \circ \sigma$ percorre S_n quando σ percorre S_n, ou seja, $S_n = \{\tau \circ \sigma : \sigma \in S_n\}$.

8.58 Seja $\sigma \in S_n$ tal que $\sigma(n) = n$. Seja $\sigma^* \in S_{n-1}$ definida por $\sigma^*(x) = \sigma(x)$.

(a) Mostre que $\sigma^* = \text{sgn}\ \sigma$,

(b) Mostre que, quando σ percorre S_n, mantendo $\sigma(n) = n$, σ^* percorre S_{n-1}; ou seja,

$$S_{n-1} = \{\sigma^* :\ \sigma \in S_n, \sigma(n) = n\}.$$

8.59 Considere uma permutação $\sigma = j_1 j_2 \ldots j_n$. Sejam $\{e_i\}$ a base canônica de K^n e A a matriz cuja i-ésima linha é e_{j_i} [ou seja, $A = (e_{j_1}, e_{j_2}, \ldots, e_{j_n})$]. Mostre que $|A| = \text{sgn}\ \sigma$.

Determinante de operador linear

8.60 Encontre o determinante de cada uma das transformações lineares seguintes.

(a) $T{:}\mathbf{R}^2 \to \mathbf{R}^2$ definida por $T(x, y) = (2x - 9y,\ \ 3x - 5y)$,

(b) $T{:}\mathbf{R}^3 \to \mathbf{R}^3$ definida por $T(x, y, z) = (3x - 2z,\ \ 5y + 7z,\ \ x + y + z)$,

(c) $T{:}\mathbf{R}^3 \to \mathbf{R}^2$ definida por $T(x, y, z) = (2x + 7y - 4z,\ \ 4x - 6y + 2z)$.

8.61 Seja $\mathbf{D}{:}V \to V$ o operador derivada, ou seja, $\mathbf{D}(f(t)) = df/dt$. Encontre det($\mathbf{D}$) sendo V o espaço vetorial de funções com a base dada. (a) $\{1, t, \ldots, t^5\}$, (b) $\{e^t, e^{2t}, e^{3t}\}$, (c) {sen t, cos t}.

8.62 Demonstre o Teorema 8.13. Sejam F e G operadores lineares de um espaço vetorial V. Então
(i) $\det(F \circ G) = \det(F)\ \det(G)$, (ii) F é invertível se, e só se, $\det(F) \neq 0$.

8.63 Demonstre (a) $\det(\mathbf{1}_V) = 1$, sendo $\mathbf{1}_V$ o operador identidade, (b) $\det(T^{-1}) = \det(T)^{-1}$, se T^{-1} for invertível.

Problemas variados

8.64 Encontre o volume $V(S)$ do paralelepípedo S de $\mathbf{R}^3$ determinado pelos vetores dados.

(a) $u_1 = (1, 2, -3)$, $u_2 = (3, 4, -1)$, $u_3 = (2, -1, 5)$,

(b) $u_1 = (1, 1, 3)$, $u_2 = (1, -2, -4)$, $u_3 = (4, 1, 5)$.

8.65 Encontre o volume $V(S)$ do paralelepípedo S de $\mathbf{R}^4$ determinado pelos vetores dados.

$$u_1 = (1, -2, 5, -1),\ \ u_2 = (2, 1, -2, 1),\ \ u_3 = (3, 0, 1 - 2),\ \ u_4 = (1, -1, 4, -1)$$

8.66 Seja V o espaço vetorial das matrizes $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ de ordem 2 sobre $\mathbf{R}$. Decida se a aplicação $D{:}V \to \mathbf{R}$ dada é bilinear (em relação às linhas).

(a) $D(M) = a + d$, (c) $D(M) = ac - bd$, (e) $D(M) = 0$,
(b) $D(M) = ad$, (d) $D(M) = ab - cd$, (f) $D(M) = 1$.

8.67 Seja A uma matriz quadrada de ordem n. Mostre que $|kA| = k^n|A|$.

8.68 Sejam A, B, C, D matrizes quadradas de ordem n que comutam. Considere a matriz quadrada de ordem $2n$ em blocos $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Prove que $|M| = |A||D| - |B||C|$. Mostre que a afirmação pode ser falsa se as matrizes não comutarem.

8.69 Seja A uma matriz ortogonal, isto é, $A^TA = I$. Mostre que $\det(A) = \pm 1$.

8.70 Seja V o espaço das matrizes quadradas de ordem m vistas como m-uplas de vetores linha. Seja $D{:}V \to K$ uma aplicação m-linear alternada. Mostre que

(a) $D(\dots, A, \dots, B, \dots) = -D(\dots, B, \dots, A, \dots)$; isto é, o sinal muda se trocarmos a posição de duas linhas;

(b) se A_1, A_2, ..., A_m são linearmente independentes, então $D(A_1, A_2, \dots, A_m) = 0$.

8.71 Sejam V o espaço das matrizes quadradas de ordem m (como no Problema 8.70) e $D{:}\ V \to K$. Mostre que a afirmação seguinte é equivalente a D ser alternada.

$$D(A_1, A_2, \dots, A_n) = 0 \qquad \text{sempre que} \qquad A_i = A_{i+1}, \text{ para algum } i.$$

Seja V o espaço das matrizes quadradas de ordem n sobre K. Seja $B \in V$ invertível, de modo que $\det(B) \neq 0$. Defina $D{:}\ V \to K$ por $D(A) = \det(AB)/\det(B)$, com $A \in V$. Logo,

$$D(A_1, A_2, \dots, A_n) = \det(A_1B, A_2B, \dots, A_nB)/\det(B)$$

em que A_i é a i-ésima linha de A e, portanto, A_iB é a i-ésima linha de AB. Mostre que D é multilinear alternada e que $D(I) = 1$. (Alguns textos usam esse método para provar que $|AB| = |A||B|$.)

8.72 Mostre que $g = g(x_1, \dots, x_n) = (-1)^n V_{n-1}(x)$, onde $g = g(x_i)$ é o produto das diferenças do Problema 8.19, $x = x_n$ e V_{n-1} é o *determinante de Vandermonde* definido por

$$V_{n-1}(x) \equiv \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ x_1 & x_2 & \dots & x_{n-1} & x \\ x_1^2 & x_2^2 & \dots & x_{n-1}^2 & x^2 \\ \dots & \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_{n-1}^{n-1} & x^{n-1} \end{vmatrix}$$

8.73 Seja A uma matriz qualquer. Mostre que o sinal de um menor $A[I, J]$ é igual ao de seu menor complementar $A[I', J']$.

8.74 Seja A uma matriz quadrada de ordem n. O *posto determinante* de A é a ordem da maior submatriz quadrada de A (obtida suprimindo linhas e colunas de A) cujo determinante é não nulo. Mostre que o posto determinante de A é igual ao posto de A, que é o maior número de linhas (ou colunas) linearmente independentes de A.

Respostas dos Problemas Complementares

Notação: $M = [R_1; \quad R_2; \quad \ldots]$ denota uma matriz de linhas R_1, R_2, ...

8.38 (a) -22, (b) -13, (c) 46, (d) -21, (e) $a^2 + ab + b^2$

8.39 (a) $3, 10$; (b) $5, -2$

8.40 (a) 21, (b) -11, (c) 100, (d) 0

8.41 (a) -131, (b) -55

8.42 (a) 33, (b) 0, (c) 45

8.43 (a) -32, (b) -14, (c) -468

8.44 (a) $|A| = -2$, adj $A = [-1, -1, 1; \quad -1, 1, -1; \quad 2, -2, 0]$,
(b) $|A| = -1$, adj $A = [1, 0, -2; \quad -3, -1, 6; \quad 2, 1, -5]$. Também, $A^{-1} = (\text{adj } A)/|A|$

8.45 (a) $[-16, -29, -26, -2; \quad -30, -38, -16, 29; \quad -8, 51, -13, -1; \quad -13, 1, 28, -18]$,
(b) $[21, -14, -17, -19; \quad -44, 11, 33, 11; \quad -29, 1, 13, 21; \quad 17, 7, -19, -18]$

8.46 (a) adj $A = [d, -b; \quad -c, a]$, (c) $A = kI$

8.49 (a) $-3, -3$, (b) $-23, -23$, (c) $3, -3$, (d) $17, -17$

8.50 (a) $-2, -17, 73$, (b) $7, 10, 105$, (c) $13, 54, 0$

8.51 (a) $-6, 13, 62, -219$; (b) $7, -37, 30, 20$

8.52 (a) $x = \frac{21}{26}, y = \frac{29}{26}$; (b) $x = -\frac{5}{13}, y = \frac{1}{13}$; (c) $x = -\frac{c}{a}, y = -\frac{c}{b}$

8.53 (a) $x = 5, y = 2, z = 1$, (b) Como $D = 0$, o sistema não pode ser resolvido com determinantes.

8.55 (a) $\text{sgn } \sigma = 1, \text{sgn } \tau = -1, \text{sgn } \pi = -1$

8.56 (a) $\tau \circ \sigma = 53142$, (b) $\pi \circ \sigma = 52413$, (c) $\sigma^{-1} = 32154$, (d) $\tau^{-1} = 14253$

8.60 (a) $\det(T) = 17$, (b) $\det(T) = 4$, (c) Não está definido.

8.61 (a) 0, (b) 6, (c) 1

8.64 (a) 18, (b) 0

8.65 17

8.66 (a) Não é, (b) É, (c) É, (d) Não é, (e) É, (f) Não é.

Diagonalização: Autovalores e Autovetores

9.1 INTRODUÇÃO

As ideias apresentadas neste capítulo podem ser discutidas em dois contextos.

Contexto de matrizes

Seja dada uma matriz quadrada A de ordem n. Dizemos que a matriz A é *diagonalizável* se existir uma matriz não singular P tal que

$$B = P\ ^{1}AP$$

seja diagonal. Neste capítulo estudamos a diagonalização de uma matriz A. Em particular, fornecemos um algoritmo para encontrar a matriz P, quando existir.

Contexto de operadores lineares

Seja dado um operador linear $T: V \to V$. Dizemos que o operador T é *diagonalizável* se existir uma base S de V tal que a representação matricial D de T em relação à base S for uma matriz diagonal. Neste capítulo estudamos condições sob as quais o operador T é diagonalizável.

Equivalência dos dois contextos

Os dois conceitos apresentados são, essencialmente, o mesmo. Mais precisamente, uma matriz quadrada A pode ser vista como o operador linear F definido por

$$F(X) = AX$$

onde X é um vetor coluna e $B = P^{-1}AP$ representa F em relação a um novo sistema de coordenadas (base) S cujos elementos são as colunas de P. Por outro lado, qualquer operador linear T pode ser representado por uma matriz A em relação a alguma base e, escolhendo alguma outra base, T é representado pela matriz

$$B = P^{-1}AP$$

em que P é a matriz de mudança de base.

A maioria dos teoremas será enunciada de duas formas, uma em termos de matrizes A e outra em termos de operadores lineares T.

Papel do corpo *K*

O corpo K subjacente não teve, até agora, papel especial algum nas nossas discussões sobre espaços vetoriais e transformações lineares. No entanto, a diagonalização de uma matriz A ou de um operador linear T depende das raízes de um polinômio $\Delta(t)$ sobre K e essas raízes dependem de K. Por exemplo, suponha que $\Delta(t) = t^2 + 1$. Então $\Delta(t)$ não tem raízes se $K = \mathbf{R}$, o corpo dos reais, mas $\Delta(t)$ tem as raízes $\pm i$ se $K = \mathbf{C}$, o corpo dos complexos. Além disso, encontrar as raízes de um polinômio de grau maior do que dois é, em geral, um assunto em si (frequentemente discutido em disciplinas de Cálculo Numérico). Por isso, nossos exemplos geralmente levam a polinômios $\Delta(t)$ cujas raízes são encontradas facilmente.

9.2 POLINÔMIOS DE MATRIZES

Considere um polinômio $f(t) = a_n t^n + \cdots + a_1 t + a_0$ sobre um corpo K. Na Seção 2.8 definimos, para matrizes quadradas A quaisquer,

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I$$

onde I é a matriz identidade. Em particular, dizemos que A é uma *raiz* de $f(t)$ se $f(A) = 0$, a matriz nula.

Exemplo 9.1 Seja $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Então $A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$. Sejam

$$f(t) = 2t^2 - 3t + 5 \qquad \text{e} \qquad g(t) = t^2 - 5t - 2$$

Então

$$f(A) = 2A^2 - 3A + 5I = \begin{bmatrix} 14 & 20 \\ 30 & 44 \end{bmatrix} + \begin{bmatrix} -3 & -6 \\ -9 & -12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 16 & 14 \\ 21 & 37 \end{bmatrix}$$

e

$$g(A) = A^2 - 5A - 2I = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} + \begin{bmatrix} -5 & -10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Assim, A é um zero de $g(t)$.

Vale o teorema seguinte (demonstrado no Problema 9.7).

Teorema 9.1 Sejam f e g polinômios. Dados qualquer matriz quadrada A e qualquer escalar k temos

(i) $(f+g)(A) = f(A) + g(A)$ (iii) $(kf)(A) = kf(A)$

(ii) $(fg)(A) = f(A)g(A)$ (iv) $f(A)g(A) = g(A)f(A)$.

Observe que (iv) nos diz que dois polinômios quaisquer de A comutam.

Matrizes e operadores lineares

Seja, agora, $T\colon V \to V$ um operador linear de um espaço vetorial V. As potências de T podem ser definidas pela composição, como segue.

$$T^2 = T \circ T, \qquad T^3 = T^2 \circ T, \qquad \ldots$$

Também, dado qualquer polinômio $f(t) = a_n t^n + \cdots + a_1 t + a_0$, definimos $f(T)$ da mesma forma que para matrizes, ou seja,

$$f(T) = a_n T^n + \cdots + a_1 T + a_0 I$$

onde, agora, I é o operador identidade. Também dizemos que T é uma *raiz* de $f(t)$ se $f(T) = 0$, o operador nulo. Observamos que as relações do Teorema 9.1 são válidas tanto para operadores lineares quanto para matrizes.

OBSERVAÇÃO Seja A uma representação matricial de um operador linear T. Então $f(A)$ é a representação matricial de $f(T)$ e, em particular, $f(T) = 0$ se, e só se, $f(A) = 0$.

9.3 POLINÔMIO CARACTERÍSTICO, TEOREMA DE CAYLEY-HAMILTON

Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n. A matriz $M = A - tI_n$, em que I_n é a matriz quadrada identidade de ordem n e t é uma incógnita, pode ser obtida subtraindo t de cada elemento diagonal de A. A matriz simétrica na soma de M é a matriz $tI_n - A$ e seu determinante

$$\Delta(t) = \det(tI_n - A) = (-1)^n \det(A - tI_n)$$

que é um polinômio de grau n em t, é denominado *polinômio característico* de A.

Enunciamos, agora, um teorema importante da Álgebra Linear (demonstrado no Problema 9.8).

Teorema 9.2 (Cayley-Hamilton) Cada matriz A é uma raiz de seu polinômio característico.

OBSERVAÇÃO Se $A = [a_{ij}]$ for uma matriz triangular, então $tI - A$ é uma matriz triangular com entradas diagonais $t - a_{ii}$ e, portanto,

$$\Delta(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

Observe que as raízes de $\Delta(t)$ são as entradas diagonais de A.

Exemplo 9.2 O polinômio característico de $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$ é

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 \\ -4 & t-5 \end{vmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

Conforme afirma o Teorema de Cayley-Hamilton, A é uma raiz de $\Delta(t)$, pois

$$\Delta(A) = A^2 - 6A - 7I = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix} + \begin{bmatrix} -6 & -18 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} -7 & 0 \\ 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Sejam, agora, A e B são matrizes semelhantes, digamos, $B = P^{-1}AP$, em que P é invertível. Mostremos que A e B têm o mesmo polinômio característico. Usando $tI = P^{-1}tIP$, temos

$$\begin{aligned} \Delta_B(t) &= \det(tI - B) = \det(tI - P^{-1}AP) = \det(P^{-1}tIP - P^{-1}AP) \\ &= \det[P^{-1}(tI - A)P] = \det(P^{-1}) \det(tI - A) \det(P) \end{aligned}$$

Como os determinantes são escalares, decorre que comutam e, como $\det(P^{-1}) \det(P) = 1$, obtemos

$$\Delta_B(t) = \det(tI - A) = \Delta_A(t)$$

Assim, demonstramos o teorema a seguir.

Teorema 9.3 Matrizes semelhantes têm o mesmo polinômio característico.

Polinômios característicos de graus 2 e 3

Há fórmulas simples para os polinômios característicos de matrizes de ordens 2 e 3.

(a) Seja $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Então

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \det(A) = t^2 - \operatorname{tr}(A)\, t + \det(A)$$

Aqui, tr(A) denota o traço de A, isto é, a soma dos elementos diagonais de A.

(b) Seja $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Então

$$\Delta(t) = t^3 - \mathrm{tr}(A)\, t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

Aqui, A_{11}, A_{22}, A_{33} denotam, respectivamente, os cofatores de a_{11}, a_{22}, a_{33}.

Exemplo 9.3 Encontre o polinômio característico de cada uma das matrizes seguintes.

(a) $A = \begin{bmatrix} 5 & 3 \\ 2 & 10 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -1 \\ 6 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 5 & -2 \\ 4 & -4 \end{bmatrix}$.

(a) Temos $\mathrm{tr}(A) = 5 + 10 = 15$ e $|A| = 50 - 6 = 44$, portanto, $\Delta(t) + t^2 - 15t + 44$.
(b) Temos $\mathrm{tr}(B) = 7 + 2 = 9$ e $|B| = 14 + 6 = 20$, portanto, $\Delta(t) = t^2 - 9t + 20$.
(c) Temos $\mathrm{tr}(C) = 5 - 4 = 1$ e $|C| = -20 + 8 = -12$, portanto, $\Delta(t) = t^2 - t - 12$.

Exemplo 9.4 Encontre o polinômio característico de $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$.

Temos $\mathrm{tr}(A) = 1 + 3 + 9 = 13$. Os cofatores dos elementos diagonais são

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 3 & 9 \end{vmatrix} = 21, \qquad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 9 \end{vmatrix} = 7, \qquad A_{33} = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3$$

Assim, $A_{11} + A_{22} + A_{33} = 31$ e, também, $|A| = 27 + 2 + 0 - 6 - 6 - 0 = 17$. Dessa forma,

$$\Delta(t) = t^3 - 13t^2 + 31t - 17$$

OBSERVAÇÃO Os coeficientes do polinômio característico $\Delta(t)$ de uma matriz quadrada de ordem 3 são os seguintes, com sinais alternados.

$$S_1 = \mathrm{tr}(A), \qquad S_2 = A_{11} + A_{22} + A_{33}, \qquad S_3 = \det(A)$$

Observamos que cada S_k é a soma de todos os menores principais de A de ordem k.

O próximo teorema, cuja demonstração está além do alcance deste livro, nos diz que esse resultado é verdadeiro em geral.

Teorema 9.4 Seja A uma matriz quadrada de ordem n. O polinômio característico de A é

$$\Delta(t) = t^n - S_1 t^{n-1} + S_2 t^{n-2} + \cdots + (-1)^n S_n$$

onde S_k é a soma de todos os menores principais de ordem k.

Polinômio característico de um operador linear

Seja, agora, $T\colon V \to V$ um operador linear de um espaço vetorial V de dimensão finita. Definimos o *polinômio característico* $\Delta(t)$ de T como sendo o polinômio característico de qualquer representação matricial de T. Lembre-se que se A e B forem representações matriciais de T, então $B = P^{-1}AP$, sendo P a matriz de mudança de base. Assim, A e B são semelhantes e, pelo Teorema 9.3, A e B têm o mesmo polinômio característico. Por isso, o polinômio característico de T independe da particular escolha da base na qual calculamos a representação matricial de T.

Como $f(T) = 0$ se, e só se, $f(A) = 0$, sendo $f(t)$ um polinômio qualquer e A uma representação matricial qualquer de T, temos o teorema seguinte análogo para operadores lineares.

Teorema 9.2′ (Cayley-Hamilton) Um operador linear T é uma raiz de seu polinômio característico.

9.4 DIAGONALIZAÇÃO, AUTOVALORES E AUTOVETORES

Seja A uma matriz quadrada de ordem n. Então A pode ser representada por (ou é semelhante a) uma matriz diagonal $D = \text{diag}(k_1, k_2, ..., k_n)$ se, e só se, existe uma base S de vetores (coluna) $u_1, u_2, ..., u_n$ tal que

$$\begin{aligned} Au_1 &= k_1u_1 \\ Au_2 &= \qquad k_2u_2 \\ &\cdots\cdots\cdots\cdots \\ Au_n &= \qquad\qquad k_nu_n \end{aligned}$$

Neste caso, dizemos que A é *diagonalizável*. Além disso, se P denotar a matriz não singular cujas colunas são, respectivamente, os vetores da base $u_1, u_2, ..., u_n$, então $D = P^{-1}AP$.

Essa observação nos leva à definição seguinte.

DEFINIÇÃO Seja A uma matriz quadrada qualquer. Dizemos que um escalar λ é um *autovalor* de A se existe um vetor (coluna) não nulo v tal que

$$Av = \lambda v$$

Qualquer vetor satisfazendo essa relação é denominado *autovetor* de A *associado* (ou *correspondente*, ou *pertencente*) ao autovalor λ.

Observe que qualquer múltiplo escalar kv de um autovetor v associado a λ também é um autovetor associado a λ, pois

$$A(kv) = k(Av) = k(\lambda v) = \lambda(kv)$$

O conjunto E_λ de todos esses autovetores é um subespaço de V (Problema 9.19), denominado *autoespaço* de λ. (Se $E_\lambda = 1$, dizemos que E_λ é uma *reta invariante* e λ é um *fator de escala*.)

Os termos *valor característico* e *vetor característico*, ou *valor próprio* e *vetor próprio*, também são utilizados em vez de autovalor e autovetor.

Essas observações e definições nos dão o teorema a seguir.

Teorema 9.5 Uma matriz quadrada A de ordem n é semelhante a uma matriz diagonal D se, e só se, A tem n autovetores linearmente independentes. Nesse caso, os elementos diagonais de D são os autovalores associados e $D = P^{-1}AP$, onde P é a matriz cujas colunas são os autovetores.

Suponha que uma matriz A possa ser diagonalizada, digamos, $P^{-1}AP = D$, com D diagonal. Então A possui a *fatoração diagonal* extremamente útil

$$A = PDP^{-1}$$

Usando essa fatoração, a álgebra de A se reduz à álgebra da matriz diagonal D, que é muito facilmente calculada. Mais precisamente, digamos que $D = \text{diag}(k_1, k_2, ..., k_n)$. Então

$$A^m = (PDP^{-1})^m = PD^mP^{-1} = P\ \text{diag}(k_1^m, \ldots, k_n^m)P^{-1}$$

Mais geralmente, dado qualquer polinômio $f(t)$, temos

$$f(A) = f(PDP^{-1}) = Pf(D)P^{-1} = P\ \text{diag}(f(k_1), f(k_2), \ldots, f(k_n))P^{-1}$$

Além disso, se os elementos diagonais de D forem não negativos, considere

$$B = P\ \text{diag}(\sqrt{k_1}, \sqrt{k_2}, \ldots, \sqrt{k_n})\ P^{-1}$$

Então B é uma *raiz quadrada não negativa* de A, ou seja, $B^2 = A$ e os autovalores de B são não negativos.

Exemplo 9.5 Sejam $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$, $v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ e $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Então

$$Av_1 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = v_1 \qquad \text{e} \qquad Av_2 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4v_2$$

Assim, v_1 e v_2 são autovetores de A associados, respectivamente, aos autovalores $\lambda_1 = 1$ e $\lambda_2 = 4$. Observe que v_1 e v_2 são linearmente independentes e, portanto, formam uma base de $\mathbf{R}^2$. Dessa forma, A é diagonalizável. Além disso, seja P a matriz cujas colunas são esses autovetores v_1 e v_2, como segue.

$$P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \text{e, portanto, } P^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Então A é semelhante à matriz diagonal

$$D = P^{-1}AP = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Como era de se esperar, os elementos diagonais 1 e 4 de D são os autovalores associados aos autovetores v_1 e v_2, respectivamente, que são as colunas de P. Em particular, A tem a fatoração

$$A = PDP^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Por isso,

$$A^4 = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 256 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 171 & 85 \\ 170 & 86 \end{bmatrix}$$

Além disso, dado $f(t) = t^3 - 5t^2 + 3t + 6$, temos $f(1) = 5$ e $f(4) = 2$ e

$$f(A) = Pf(D)P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}$$

Finalmente, calculemos a "raiz quadrada positiva" de A. Para isso, usamos $\sqrt{1} = 1$ e $\sqrt{4} = 2$ para obter a matriz

$$B = P\sqrt{D}P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{4}{3} \end{bmatrix}$$

tal que $B^2 = A$, sendo que B tem os autovalores positivos 1 e 2.

OBSERVAÇÃO Em todo este capítulo, usamos o fato seguinte.

$$\text{Se } P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ então } P^{-1} = \begin{bmatrix} d/|P| & -b/|P| \\ -c/|P| & a/|P| \end{bmatrix}$$

Ou seja, para obter P^{-1}, trocamos a posição dos elementos diagonais a e d de P, tomamos o simétrico dos elementos não diagonais b e c e dividimos cada entrada pelo determinante $|P|$.

Propriedades dos autovalores e autovetores

No Exemplo 9.5 vimos as vantagens de uma representação (fatoração) diagonal de uma matriz quadrada. No teorema seguinte (demonstrado no Problema 9.20), listamos as propriedades que nos permitem encontrar uma tal representação.

Teorema 9.6 Seja A uma matriz quadrada. As afirmações dadas são equivalentes.

(i) O escalar λ é um autovalor de A.
(ii) A matriz $M = A - \lambda I$ é singular.
(iii) O escalar λ é uma raiz do polinômio característico $\Delta(t)$ de A.

O autoespaço E_λ de um autovalor λ é o espaço solução do sistema homogêneo $MX = 0$, onde $M = A - \lambda I$, ou seja, M é obtida subtraindo λ de cada elemento diagonal de A.

Algumas matrizes não possuem autovalores e, consequentemente, autovetores. No entanto, usando o Teorema 9.6 e o Teorema Fundamental da Álgebra (todo polinômio sobre o corpo dos complexos **C** tem alguma raiz), obtemos o resultado seguinte.

Teorema 9.7 Seja A uma matriz quadrada sobre o corpo dos complexos **C**. Então A possui pelo menos um autovalor.

Os teoremas a seguir serão utilizados mais tarde. (O teorema equivalente ao Teorema 9.8 para operadores lineares está demonstrado no Problema 9.21 e o Teorema 9.9 está demonstrado no Problema 9.22.)

Teorema 9.8 Suponha que $v_1, v_2, ..., v_n$ sejam autovetores de uma matriz A associados a autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Então $v_1, v_2, ..., v_n$ são linearmente independentes.

Teorema 9.9 Suponha que o polinômio característico $\Delta(t)$ de uma matriz quadrada A de ordem n seja um produto de n fatores distintos, digamos, $\Delta(t) = (t - a_1)(t - a_2)\cdots(t - a_n)$. Então A é semelhante à matriz diagonal $D = \text{diag}(a_1, a_2, ..., a_n)$.

Se λ é um autovalor de uma matriz A, então a *multiplicidade algébrica* de λ é definida como a multiplicidade de λ como raiz do polinômio característico de A e a *multiplicidade geométrica* de λ é a dimensão do autoespaço, dim E_λ. Vale o teorema seguinte (cujo equivalente para operadores lineares está demonstrado no Problema 9.23).

Teorema 9.10 A multiplicidade geométrica de um autovalor λ de uma matriz A não é maior do que a multiplicidade algébrica.

Diagonalização de operadores lineares

Considere um operador linear $T: V \to V$. Dizemos que T é *diagonalizável* se T puder ser representado por uma matriz diagonal D. Assim, T é diagonalizável se, e só se, existir uma base $S = \{u_1, u_2, ..., u_n\}$ de V tal que

$$\begin{aligned} T(u_1) &= k_1u_1 \\ T(u_2) &= \quad k_2u_2 \\ &\cdots\cdots\cdots\cdots \\ T(u_n) &= \qquad k_nu_n \end{aligned}$$

Nesse caso, T é representado pela matriz diagonal

$$D = \text{diag}(k_1, k_2, ..., k_n)$$

em relação à base S.

Essa observação nos leva às definições e teoremas seguintes, análogos às definições e teoremas para matrizes que já apresentamos.

DEFINIÇÃO Seja T um operador linear. Dizemos que um escalar λ é um *autovalor* de T se existe um vetor não nulo v tal que $T(v) = \lambda v$.

Qualquer vetor satisfazendo essa relação é denominado *autovetor* de T *associado* (ou *correspondente*, ou *pertencente*) ao autovalor λ.

O conjunto E_λ de todos os autovetores associados a um autovalor λ é um subespaço de V, denominado *autoespaço* de λ. (Alternativamente, λ é um autovalor de T se $\lambda I - T$ é singular e, nesse caso, E_λ é o núcleo de $\lambda I - T$.) As *multiplicidades algébrica* e *geométrica* de um autovalor λ de um operador T são definidas da mesma maneira que as de um autovalor de uma matriz A.

Os teoremas seguintes são válidos para operadores lineares T de espaços vetoriais V de dimensão finita.

Teorema 9.5′ T pode ser representado por uma matriz diagonal D se, e só se, existe uma base S de V consistindo em autovetores de T. Nesse caso, os elementos diagonais de D são os autovalores associados.

Teorema 9.6′ Seja T um operador linear. As afirmações dadas são equivalentes.

(i) O escalar λ é um autovalor de T.
(ii) O operador $\lambda I - T$ é singular.
(iii) O escalar λ é uma raiz do polinômio característico $\Delta(t)$ de T.

Teorema 9.7′ Se V é um espaço vetorial complexo, então T possui pelo menos um autovalor.

Teorema 9.8′ Suponha que $v_1, v_2, ..., v_n$ sejam autovetores de um operador linear T associados a autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Então $v_1, v_2, ..., v_n$ são linearmente independentes.

Teorema 9.9′ Suponha que o polinômio característico $\Delta(t)$ de T seja um produto de n fatores distintos, digamos, $\Delta(t) = (t - a_1)(t - a_2) \cdots (t - a_n)$. Então T pode ser representado pela matriz diagonal $D = \text{diag}(a_1, a_2, ..., a_n)$.

Teorema 9.10′ A multiplicidade geométrica de um autovalor λ de T não é maior do que a multiplicidade algébrica.

OBSERVAÇÃO O teorema a seguir reduz a procura de uma diagonalização de um operador linear T à diagonalização de uma matriz A.

Teorema 9.11 Seja A uma representação matricial de T. Então T é diagonalizável se, e só se, A é diagonalizável.

9.5 CÁLCULO DE AUTOVALORES E AUTOVETORES, DIAGONALIZAÇÃO DE MATRIZES

Nesta seção oferecemos um algoritmo para calcular autovalores e autovetores de uma dada matriz quadrada A dada e para decidir se existe ou não uma matriz não singular P tal que $P^{-1}AP$ seja diagonal.

Algoritmo 9.1 (Algoritmo de Diagonalização) É dada uma matriz quadrada A de ordem n.

Passo 1 Encontre o polinômio característico $\Delta(t)$ de A.

Passo 2 Encontre as raízes de $\Delta(t)$ para obter os autovalores de A.

Passo 3 Repita (a) e (b) para cada autovalor λ de A.

(a) Forme a matriz $M = A - \lambda I$ subtraindo λ de cada elemento diagonal de A.
(b) Encontre uma base para o espaço solução do sistema homogêneo $MX = 0$. (Esses vetores da base são autovetores linearmente independentes de A associados a λ.)

Passo 4 Considere a coleção $S = \{v_1, v_2, ..., v_m\}$ de todos os autovetores obtidos no Passo 3.

(a) Se $m \neq n$, então A não é diagonalizável.

(b) Se $m = n$, então A é diagonalizável. Mais precisamente, seja P a matriz cujas colunas são os autovetores $v_1, v_2, ..., v_m$. Então

$$D = P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

onde λ_i é o autovalor associado ao autovetor v_i.

Exemplo 9.6 Aplique o algoritmo de diagonalização à matriz $A = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}$.

(1) Calculamos o polinômio característico $\Delta(t)$ de A. Temos

$$\operatorname{tr}(A) = 4 - 1 = -3, \qquad |A| = -4 - 6 = -10;$$

logo,

$$\Delta(t) = t^2 - 3t - 10 = (t-5)(t+2)$$

(2) Tomamos $\Delta(t) = (t-5)(t+2) = 0$. As raízes $\lambda_1 = 5$ e $\lambda_2 = -2$ são os autovalores de A.

(3) (i) Encontramos um autovetor v_1 de A associado ao autovalor $\lambda_1 = 5$ subtraindo $\lambda_1 = 5$ de cada elemento diagonal de A e obtendo a matriz $M = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$. Os autovetores associados a $\lambda_1 = 5$ formam o espaço solução do sistema homogêneo $MX = 0$, isto é,

$$\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{ou} \quad \begin{matrix} -x + 2y = 0 \\ 3x - 6y = 0 \end{matrix} \quad \text{ou} \quad -x + 2y = 0$$

O sistema tem apenas uma variável livre. Assim, uma solução não nula, por exemplo, $v_1 = (2, 1)$, é um autovetor que gera o autoespaço de $\lambda_1 = 5$.

(ii) Encontramos um autovetor v_2 de A associado ao autovalor $\lambda_2 = -2$ subtraindo -2 de cada elemento diagonal de A (ou, então, somando 2) e obtendo a matriz

$$M = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \text{e o sistema homogêneo} \quad \begin{matrix} 6x + 2y = 0 \\ 3x + \ \ y = 0 \end{matrix} \quad \text{ou} \quad 3x + y = 0.$$

O sistema só tem uma solução independente. Assim, uma solução não nula, digamos, $v_2 = (-1, 3)$, é um autovetor que gera o autoespaço de $\lambda_2 = -2$.

(4) Seja P a matriz cujas colunas são os autovetores v_1 e v_2. Então

$$P = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad \text{e, portanto,} \quad P^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}$$

Dessa forma, $D = P^{-1}AP$ é a matriz diagonal cujas entradas diagonais são os autovalores associados, isto é,

$$D = P^{-1}AP = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix}$$

Exemplo 9.7 Considere a matriz $B = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$. Temos

$$\operatorname{tr}(B) = 5 + 3 = 8, \qquad |B| = 15 + 1 = 16; \qquad \text{portanto,} \qquad \Delta(t) = t^2 - 8t + 16 = (t-4)^2$$

Dessa forma, $\lambda = 4$ é o único autovalor de B.

Subtraindo $\lambda = 4$ de cada elemento diagonal de B, obtemos a matriz

$$M = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \text{ e o sistema homogêneo } \quad \begin{matrix} x - y = 0 \\ x - y = 0 \end{matrix} \quad \text{ou} \quad x - y = 0$$

O sistema tem uma única solução independente, por exemplo $x = 1$, $y = 1$. Assim, $v = (1, 1)$ e seus múltiplos são os únicos autovetores de B. Dessa forma, B não é diagonalizável, porque não existe uma base consistindo em autovetores de B.

Exemplo 9.8 Considere a matriz $A = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix}$. Aqui, $\text{tr}(A) = 3 - 3 = 0$ e $|A| = -9 + 10 = 1$. Assim, $\Delta(t) = t^2 + 1$ é o polinômio característico de A. Consideramos os dois casos a seguir.

(a) Se A é uma matriz sobre o corpo **R** dos reais, então $\Delta(t)$ não tem raízes (reais). Assim, A não possui autovalores nem autovalores e, portanto, A não é diagonalizável.

(b) Se A é uma matriz sobre o corpo **C** dos complexos, então $\Delta(t) = (t - i)(t + i)$ tem duas raízes, i e $-i$. Assim, A tem os dois autovalores distintos i e $-i$ e, portanto, A possui dois autovetores linearmente independentes. Dessa forma, existe uma matriz não singular P sobre o corpo **C** dos complexos tal que

$$P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

Assim, A é diagonalizável (sobre **C**).

9.6 DIAGONALIZAÇÃO DE MATRIZES REAIS SIMÉTRICAS E FORMAS QUADRÁTICAS

Existem muitas matrizes reais que não são diagonalizáveis. De fato, algumas matrizes reais podem sequer possuir algum autovalor (real). Contudo, com matrizes reais simétricas, esses problemas não ocorrem. A saber, temos os teoremas seguintes.

Teorema 9.12 Se A é uma matriz real simétrica, então cada raiz λ do polinômio característico de A é real.

Teorema 9.13 Sejam A uma matriz real simétrica e u e v autovetores de A associados a autovalores distintos λ_1 e λ_2. Então u e v são ortogonais, isto é, $\langle u, v \rangle = 0$.

Esses dois teoremas fornecem o resultado fundamental seguinte.

Teorema 9.14 Se A é uma matriz real simétrica, então existe uma matriz ortogonal P tal que $D = P^{-1}AP$ é diagonal.

A matriz ortogonal P é obtida normalizando uma base de autovetores ortogonais de A, conforme exemplo a seguir. Nesse caso, dizemos que A é "ortogonalmente diagonalizável".

Exemplo 9.9 Seja $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$, uma matriz real simétrica. Encontre uma matriz ortogonal P tal que $P^{-1}AP$ seja diagonal.

Inicialmente, calculamos o polinômio característico $\Delta(t)$ de A. Temos

$$\text{tr}(A) = 2 + 5 = 7, \qquad |A| = 10 - 4 = 6; \quad \text{portanto,} \quad \Delta(t) = t^2 - 7t + 6 = (t - 6)(t - 1)$$

Segue que $\lambda_1 = 6$ e $\lambda_2 = 1$ são os autovalores de A.

(a) Subtraindo $\lambda_1 = 6$ de cada elemento diagonal de A obtemos a matriz

$$M = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} \text{ e o sistema homogêneo } \quad \begin{matrix} -4x - 2y = 0 \\ -2x - y = 0 \end{matrix} \quad \text{ou} \quad 2x + y = 0$$

Uma solução não nula é $u_1 = (1, -2)$.

(b) Subtraindo $\lambda_2 = 1$ de cada elemento diagonal de A obtemos a matriz

$$M = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \quad \text{e o sistema homogêneo} \quad x - 2y = 0$$

(A segunda equação foi eliminada, por ser um múltiplo da primeira equação.) Uma solução não nula é $u_2 = (2, 1)$.

Conforme afirma o Teorema 9.13, u_1 e u_2 são ortogonais. Normalizando u_1 e u_2, obtemos os vetores ortonormais

$$\hat{u}_1 = (1/\sqrt{5}, -2/\sqrt{5}) \qquad \text{e} \qquad \hat{u}_2 = (2/\sqrt{5}, 1/\sqrt{5})$$

Finalmente, seja P a matriz cujas colunas são $\hat{u}_1$ e $\hat{u}_2$, respectivamente. Então

$$P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \qquad \text{e} \qquad P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

Como era de se esperar, as entradas diagonais de $P^{-1}AP$ são os autovalores associados às colunas de P.

O procedimento desse exemplo pode ser formalizado no algoritmo seguinte, que fornece uma matriz ortogonal P tal que $P^{-1}AP$ é diagonal.

Algoritmo 9.2 (Algoritmo de Diagonalização Ortogonal) É dada uma matriz real simétrica A.

Passo 1 Encontre o polinômio característico $\Delta(t)$ de A.
Passo 2 Encontre os autovalores de A, que são as raízes de $\Delta(t)$.
Passo 3 Para cada autovalor λ de A do Passo 2, encontre uma base ortogonal do autoespaço associado.
Passo 4 Normalize todos autovetores do Passo 3, que formam uma base ortonormal de $\mathbf{R}^n$.
Passo 5 Seja P a matriz cujas colunas são os autovetores normalizados do Passo 4.

Aplicação às formas quadráticas

Seja q um polinômio real nas variáveis $x_1, x_2, ..., x_n$ tal que cada parcela de q é de grau dois, isto é,

$$q(x_1, x_2, \ldots, x_n) = \sum_i c_i x_i^2 + \sum_{i<j} d_{ij} x_i x_j, \qquad \text{com} \qquad c_i, d_{ij} \in \mathbf{R}$$

Dizemos, então, que q é uma *forma quadrática*. Se não existirem parcelas mistas da forma $x_i x_j$, isto é, se todos os d_{ij} forem nulos, dizemos que q é *diagonal*.

Toda forma quadrática q define uma matriz real simétrica $A = [a_{ij}]$, com $a_{ii} = c_i$ e $a_{ij} = a_{ji} = \frac{1}{2} d_{ij}$. A saber, q pode ser escrita no formato matricial

$$q(X) = X^T A X$$

em que $X = [x_1, x_2, ..., x_n]^T$ é o vetor coluna das variáveis. Além disso, seja $X = PY$ uma substituição linear de variáveis. Substituindo na forma quadrática, obtemos

$$q(Y) = (PY)^T A(PY) = Y^T (P^T AP) Y$$

Assim, $P^T AP$ é a representação matricial de q nas novas variáveis.

Procuramos uma matriz ortogonal P tal que a *substituição ortogonal* $X = PY$ forneça uma forma quadrática diagonal, ou seja, para a qual $P^T AP$ seja diagonal. Sendo P ortogonal, $P^T = P^{-1}$ e, portanto, $P^T AP = P^{-1}AP$. A teoria desenvolvida nesta seção fornece uma tal matriz ortogonal P.

Exemplo 9.10 Considere a forma quadrática

$$q(x,y) = 2x^2 - 4xy + 5y^2 = X^TAX, \quad \text{com} \quad A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \quad \text{e} \quad X = \begin{bmatrix} x \\ y \end{bmatrix}$$

Pelo Exemplo 9.9,

$$P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} = P^TAP, \quad \text{com} \quad P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Seja $Y = [s,t]^T$. Então a matriz P corresponde à substituição linear ortogonal $X = PY$ das variáveis x e y em termos das variáveis s e t, como segue.

$$x = \frac{1}{\sqrt{5}}s + \frac{2}{\sqrt{5}}t, \qquad y = -\frac{2}{\sqrt{5}}s + \frac{1}{\sqrt{5}}t$$

Essa substituição em $q(x, y)$ nos dá a forma quadrática diagonal $q(s,t) = 6s^2 + t^2$.

9.7 POLINÔMIO MÍNIMO

Seja A uma matriz quadrada qualquer. Seja $J(A)$ o conjunto de todos os polinômios $f(t)$ para os quais A é uma raiz, isto é, para os quais $f(A) = 0$. O conjunto $J(A)$ não é vazio em virtude do Teorema 9.2 de Cayley-Hamilton, que nos diz que o polinômio característico $\Delta_A(t)$ de A pertence a $J(A)$. Seja $m(t)$ o polinômio mônico de menor grau em $J(A)$. (Um tal polinômio $m(t)$ existe e é único.) Dizemos que $m(t)$ é o *polinômio mínimo* da matriz A.

OBSERVAÇÃO Um polinômio $f(t) \neq 0$ é dito *mônico* se seu coeficiente dominante (o do termo de maior grau) for igual a um.

Vale o teorema seguinte (demonstrado no Problema 9.33).

Teorema 9.15 O polinômio mínimo $m(t)$ de uma matriz (operador linear) A divide qualquer polinômio que tenha A como raiz. Em particular, $m(t)$ divide o polinômio característico $\Delta(t)$ de A.

Há até uma relação ainda mais forte entre $m(t)$ e $\Delta(t)$.

Teorema 9.16 O polinômio característico $\Delta(t)$ e o polinômio mínimo $m(t)$ de uma matriz A têm os mesmos fatores irredutíveis.

Esse teorema (demonstrado no Problema 9.35) não diz que $m(t) = \Delta(t)$, mas tão somente que qualquer fator irredutível de um deles necessariamente divide o outro. Em particular, como um fator linear é irredutível, $m(t)$ e $\Delta(t)$ possuem os mesmos fatores lineares. Logo, possuem as mesmas raízes. Assim, temos o teorema a seguir.

Teorema 9.17 Um escalar λ é um autovalor de uma matriz A se, e só se, λ é uma raiz do polinômio mínimo $m(t)$ de A.

Exemplo 9.11 Encontre o polinômio mínimo $m(t)$ de $A = \begin{bmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{bmatrix}$.

Primeiro calculamos o polinômio característico $\Delta(t)$ de A. Temos

$$\text{tr}(A) = 5, \qquad A_{11} + A_{22} + A_{33} = 2 - 3 + 8 = 7, \qquad \text{e} \qquad |A| = 3$$

Logo,

$$\Delta(t) = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$$

O polinômio mínimo $m(t)$ deve dividir $\Delta(t)$. Também, cada fator irredutível de $\Delta(t)$ (ou seja, $t-1$ e $t-3$) deve ser um fator de $m(t)$. Assim, $m(t)$ é, exatamente, um dos polinômios seguintes.

$$f(t) = (t-3)(t-1) \qquad \text{ou} \qquad g(t) = (t-3)(t-1)^2$$

Sabemos, pelo teorema de Cayley-Hamilton, que $g(A) = \Delta(A) = 0$. Então, resta testar $f(t)$. Temos

$$f(A) = (A-I)(A-3I) = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, $f(t) = m(t) = (t-1)(t-3) = t^2 - 4t + 3$ é o polinômio mínimo de A.

Exemplo 9.12

(a) Considere as duas matrizes quadradas de ordem r seguintes, com $a \neq 0$.

$$J(\lambda, r) = \begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{bmatrix} \qquad \text{e} \qquad A = \begin{bmatrix} \lambda & a & 0 & \ldots & 0 & 0 \\ 0 & \lambda & a & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & \lambda & a \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{bmatrix}$$

A primeira matriz, denominada Bloco de Jordan, tem λ's em cada entrada diagonal, 1 em cada entrada *superdiagonal* (que consiste nas entradas acima da diagonal) e zeros nas demais entradas. A segunda matriz A tem λ's em cada entrada diagonal, a em cada entrada superdiagonal e zeros nas demais. [Ou seja, A é uma generalização de $J(\lambda, r)$.] Mostra-se que

$$f(t) = (t-\lambda)^r$$

é o polinômio característico e também o polinômio mínimo de ambas A e $J(\lambda, r)$.

(b) Considere um polinômio mônico arbitrário

$$f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Seja $C(f)$ a matriz quadrada de ordem n com 1 em cada entrada *subdiagonal* (que consiste nas entradas abaixo da diagonal), os simétricos dos coeficientes na última coluna e zeros nas demais entradas, como segue.

$$C(f) = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}$$

Dizemos que $C(f)$ é a *matriz companheira* do polinômio $f(t)$. Mostra-se que o polinômio mínimo $m(t)$ e também o polinômio característico $\Delta(t)$ da matriz companheira $C(f)$ são iguais ao polinômio original $f(t)$.

Polinômio mínimo de um operador linear

O *polinômio mínimo* $m(t)$ de um operador linear T é definido como o polinômio mônico de menor grau para o qual T seja uma raiz. No entanto, dado qualquer polinômio $f(t)$, temos

$$f(T) = 0 \qquad \text{se, e só se,} \qquad f(A) = 0$$

onde A é uma representação matricial qualquer de T. Dessa forma, T e A têm os mesmos polinômios mínimos. Assim, os teoremas relativos a polinômios mínimos de matrizes também se aplicam a polinômios mínimos de operadores lineares. Ou seja, valem os teoremas seguintes.

Teorema 9.15′ O polinômio mínimo $m(t)$ de um operador linear T divide qualquer polinômio que tenha T como raiz. Em particular, $m(t)$ divide o polinômio característico $\Delta(t)$ de T.

Teorema 9.16′ O polinômio característico e o polinômio mínimo de um operador linear T têm os mesmos fatores irredutíveis.

Teorema 9.17′ Um escalar λ é um autovalor de um operador linear T se, e só se, λ é uma raiz do polinômio mínimo $m(t)$ de T.

9.8 POLINÔMIOS CARACTERÍSTICO E MÍNIMO DE MATRIZES EM BLOCOS

Nesta seção discutimos as relações entre certas matrizes (quadradas) em blocos e os polinômios mínimo e característico.

Polinômio característico e matrizes triangulares em blocos

Seja M uma matriz triangular em blocos, digamos, $M = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$, em que A_1 e A_2 são matrizes quadradas. Então $tI - M$ também é uma matriz triangular em blocos, com blocos diagonais $tI - A_1$ e $tI - A_2$. Assim,

$$|tI - M| = \begin{vmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{vmatrix} = |tI - A_1||tI - A_2|$$

Ou seja, o polinômio característico de M é o produto dos polinômios característicos dos blocos diagonais A_1 e A_2.

Por indução, obtemos o resultado útil a seguir.

Teorema 9.18 Seja M uma matriz triangular em blocos com blocos diagonais $A_1, A_2, \ldots, A_r$. Então o polinômio característico de M é o produto dos polinômios característicos dos blocos diagonais A_i, ou seja,

$$\Delta_M(t) = \Delta_{A_1}(t)\Delta_{A_2}(t)\ldots\Delta_{A_r}(t)$$

Exemplo 9.13 Considere a matriz $M = \left[\begin{array}{cc|cc} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ \hline 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{array}\right]$.

Então M é uma matriz triangular em blocos com blocos diagonais $A = \begin{bmatrix} 9 & -1 \\ 8 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 6 \\ -1 & 8 \end{bmatrix}$. Temos

$$\begin{aligned} \operatorname{tr}(A) = 9 + 3 = 12, \quad \det(A) = 27 + 8 = 35 \text{ e, portanto, } \Delta_A(t) = t^2 - 12t + 35 = (t-5)(t-7) \\ \operatorname{tr}(B) = 3 + 8 = 11, \quad \det(B) = 24 + 6 = 30 \text{ e, portanto, } \Delta_B(t) = t^2 - 11t + 30 = (t-5)(t-6) \end{aligned}$$

Dessa forma, o polinômio característico de M é o produto

$$\Delta_M(t) = \Delta_A(t)\Delta_B(t) = (t-5)^2(t-6)(t-7)$$

Polinômio mínimo e matrizes diagonais em blocos

Vale o teorema seguinte (demonstrado no Problema 9.36).

Teorema 9.19 Seja M uma matriz diagonal em blocos com blocos diagonais $A_1, A_2, \ldots, A_r$. Então o polinômio mínimo de M é igual ao mínimo múltiplo comum (MMC) dos polinômios mínimos dos blocos diagonais A_i.

OBSERVAÇÃO Enfatizamos que esse teorema se aplica a matrizes diagonais em blocos, ao passo que o Teorema 9.18 análogo para o polinômio característico se aplica a matrizes triangulares em blocos.

Exemplo 9.14 Encontre o polinômio característico $\Delta(t)$ e o polinômio mínimo $m(t)$ da matriz diagonal em blocos

$$M = \begin{bmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix} = \operatorname{diag}(A_1, A_2, A_3), \text{com } A_1 = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} 4 & 2 \\ 3 & 5 \end{bmatrix}, A_3 = [7]$$

Então $\Delta(t)$ é o produto dos polinômios característicos $\Delta_1(t)$, $\Delta_2(t)$, $\Delta_3(t)$ de A_1, A_2, A_3, respectivamente. Pode-se mostrar que

$$\Delta_1(t) = (t-2)^2, \qquad \Delta_2(t) = (t-2)(t-7), \qquad \Delta_3(t) = t-7$$

Assim, $\Delta(t) = (t-2)^3(t-7)^2$. [Conforme poderíamos prever, o grau de $\Delta(t)$ é 5.]

Os polinômios mínimos $m_1(t)$, $m_2(t)$, $m_3(t)$ dos blocos diagonais A_1, A_2, A_3, respectivamente, são iguais aos polinômios característicos, isto é,

$$m_1(t) = (t-2)^2, \qquad m_2(t) = (t-2)(t-7), \qquad m_3(t) = t-7$$

Mas $m(t)$ é igual ao mínimo múltiplo comum de $m_1(t)$, $m_2(t)$, $m_3(t)$. Assim, $m(t) = (t-2)^2(t-7)$.

Problemas Resolvidos

Polinômios de matrizes, polinômios característicos

9.1 Seja $A = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}$. Encontre $f(A)$, com

$$\text{(a)} \quad f(t) = t^2 - 3t + 7, \qquad \text{(b)} \quad f(t) = t^2 - 6t + 13$$

Para começar, calculamos $A^2 = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix}$. Então

(a) $f(A) = A^2 - 3A + 7I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -3 & 6 \\ -12 & -15 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -3 & -6 \\ 12 & 9 \end{bmatrix}$

(b) $f(A) = A^2 - 6A + 13I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -6 & 12 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

[Assim, A é uma raiz de $f(t)$.]

9.2 Encontre o polinômio característico $\Delta(t)$ de cada uma das matrizes dadas.

(a) $A = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -3 \\ 5 & -2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 3 & -2 \\ 9 & -3 \end{bmatrix}$

Usamos a fórmula $\Delta(t) = t^2 - \operatorname{tr}(M)\, t + |M|$ de matrizes quadradas M de ordem 2.

(a) $\operatorname{tr}(A) = 2 + 1 = 3$, $\quad |A| = 2 - 20 = -18$, $\quad$ logo, $\quad \Delta(t) = t^2 - 3t - 18$

(b) $\operatorname{tr}(B) = 7 - 2 = 5$, $\quad |B| = -14 + 15 = 1$, $\quad$ logo, $\quad \Delta(t) = t^2 - 5t + 1$

(c) $\operatorname{tr}(C) = 3 - 3 = 0$, $\quad |C| = -9 + 18 = 9$, $\quad$ logo, $\quad \Delta(t) = t^2 + 9$

9.3 Encontre o polinômio característico $\Delta(t)$ de cada uma das matrizes dadas.

(a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 4 \\ 6 & 4 & 5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 6 & -2 \\ -3 & 2 & 0 \\ 0 & 3 & -4 \end{bmatrix}$

Usamos a fórmula $\Delta(t) = t^3 - \text{tr}(A)t^2 + (A_{11} + A_{22} + A_{33})t - |A|$, em que A_{ii} é o cofator de a_{ii} da matriz quadrada $A = [a_{ij}]$ de ordem 3.

(a) $\text{tr}(A) = 1 + 0 + 5 = 6$,

$$A_{11} = \begin{vmatrix} 0 & 4 \\ 4 & 5 \end{vmatrix} = -16, \qquad A_{22} = \begin{vmatrix} 1 & 3 \\ 6 & 5 \end{vmatrix} = -13, \qquad A_{33} = \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = -6$$

$$A_{11} + A_{22} + A_{33} = -35, \qquad \text{e} \qquad |A| = 48 + 36 - 16 - 30 = 38$$

Assim,
$$\Delta(t) = t^3 - 6t^2 - 35t - 38$$

(b) $\text{tr}(B) = 1 + 2 - 4 = -1$

$$B_{11} = \begin{vmatrix} 2 & 0 \\ 3 & -4 \end{vmatrix} = -8, \qquad B_{22} = \begin{vmatrix} 1 & -2 \\ 0 & -4 \end{vmatrix} = -4, \qquad B_{33} = \begin{vmatrix} 1 & 6 \\ -3 & 2 \end{vmatrix} = 20$$

$$B_{11} + B_{22} + B_{33} = 8, \qquad \text{e} \qquad |B| = -8 + 18 - 72 = -62$$

Assim,
$$\Delta(t) = t^3 + t^2 - 8t + 62$$

9.4 Encontre o polinômio característico $\Delta(t)$ de cada uma das matrizes dadas.

$$\text{(a)}\ A = \begin{bmatrix} 2 & 5 & 1 & 1 \\ 1 & 4 & 2 & 2 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 2 & 3 \end{bmatrix}, \ \text{(b)} \quad B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 3 & 4 \\ 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

(a) A é triangular em blocos com blocos diagonais

$$A_1 = \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix} \qquad \text{e} \qquad A_2 = \begin{bmatrix} 6 & -5 \\ 2 & 3 \end{bmatrix}$$

Assim,
$$\Delta(t) = \Delta_{A_1}(t)\Delta_{A_2}(t) = (t^2 - 6t + 3)(t^2 - 9t + 28)$$

(b) Como B é triangular, $\Delta(t) = (t-1)(t-3)(t-5)(t-6)$.

9.5 Encontre o polinômio característico $\Delta(t)$ de cada um dos operadores lineares dados.

(a) $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ definido por $F(x, y) = (3x + 5y, \quad 2x - 7y)$.

(b) $\mathbf{D}\colon V \to V$ definido por $\mathbf{D}(f) = df/dt$, sendo V o espaço de funções com base $S = \{\text{sen}\ t, \cos t\}$.

O polinômio característico $\Delta(t)$ de um operador linear é igual ao polinômio característico de qualquer matriz A que represente o operador.

(a) Utilizamos a matriz A que representa T em relação à base canônica de $\mathbf{R}^2$. Temos

$$A = \begin{bmatrix} 3 & 5 \\ 2 & -7 \end{bmatrix}, \qquad \text{logo,} \qquad \Delta(t) = t^2 - \text{tr}(A)\ t + |A| = t^2 + 4t - 31$$

(b) Utilizamos a matriz A que representa o operador derivada $\mathbf{D}$ em relação à base S. Temos

$$\begin{aligned} \mathbf{D}(\text{sen}\,t) &= \cos t \quad = 0(\text{sen}\,t) + 1(\cos t) \\ \mathbf{D}(\cos t) &= -\text{sen}\,t = -1(\text{sen}\,t) + 0(\cos t) \end{aligned} \qquad \text{logo} \qquad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Assim,
$$\Delta(t) = t^2 - \text{tr}(A)\ t + |A| = t^2 + 1$$

9.6 Mostre que uma matriz A e sua transposta A^T têm o mesmo polinômio característico.

Usando a transposta, $(tI - A)^T = tI^T - A^T = tI - A^T$. Como uma matriz e sua transposta têm o mesmo determinante,

$$\Delta_A(t) = |tI - A| = |(tI - A)^T| = |tI - A^T| = \Delta_{A^T}(t)$$

9.7 Demonstre o Teorema 9.1. Sejam f e g polinômios. Dados quaisquer matriz quadrada A e escalar k,

(i) $(f+g)(A) = f(A) + g(A)$, (iii) $(kf)(A) = kf(A)$,

(ii) $(fg)(A) = f(A)g(A)$, (iv) $f(A)g(A) = g(A)f(A)$.

Sejam $f = a_n t^n + \cdots + a_1 t + a_0$ e $g = b_m t^m + \cdots + b_1 t + b_0$. Então, por definição,

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I \qquad \text{e} \qquad g(A) = b_m A^m + \cdots + b_1 A + b_0 I$$

(i) Suponha que $m \leq n$ e seja $b_i = 0$, com $i > m$. Então

$$f + g = (a_n + b_n)t^n + \cdots + (a_1 + b_1)t + (a_0 + b_0)$$

Logo,

$$\begin{aligned}(f+g)(A) &= (a_n + b_n)A^n + \cdots + (a_1 + b_1)A + (a_0 + b_0)I \\ &= a_n A^n + b_n A^n + \cdots + a_1 A + b_1 A + a_0 I + b_0 I = f(A) + g(A)\end{aligned}$$

(ii) Por definição, $fg = c_{n+m} t^{n+m} + \cdots + c_1 t + c_0 = \sum_{k=0}^{n+m} c_k t^k$, onde

$$c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}$$

Logo, $(fg)(A) = \sum_{k=0}^{n+m} c_k A^k$ e

$$f(A)g(A) = \left(\sum_{i=0}^{n} a_i A^i\right)\left(\sum_{j=0}^{m} b_j A^j\right) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j A^{i+j} = \sum_{k=0}^{n+m} c_k A^k = (fg)(A)$$

(iii) Por definição, $kf = ka_n t^n + \cdots + ka_1 t + ka_0$ e, portanto,

$$(kf)(A) = ka_n A^n + \cdots + ka_1 A + ka_0 I = k(a_n A^n + \cdots + a_1 A + a_0 I) = kf(A)$$

(iv) Por (ii), $g(A)f(A) = (gf)(A) = (fg)(A) = f(A)g(A)$.

9.8 Demonstre o Teorema 9.2 (Cayley-Hamilton). Cada matriz A é uma raiz de seu polinômio característico.

Sejam A uma matriz quadrada de ordem n qualquer e $\Delta(t)$ seu polinômio característico, digamos,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Denotemos por $B(t)$ a adjunta clássica da matriz $tI - A$. Os elementos de $B(t)$ são os cofatores da matriz $tI - A$ e, portanto, polinômios em t de grau, no máximo, $n - 1$. Assim,

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

onde os B_i denotam matrizes quadradas de ordem n sobre K que não dependem de t. Pela propriedade fundamental da adjunta clássica (Teorema 8.9), temos $(tI - A)B(t) = |tI - A|I$, ou

$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Removendo os parênteses e igualando as potências correspondentes de t, obtemos

$$B_{n-1} = I, \qquad B_{n-2} - AB_{n-1} = a_{n-1}I, \qquad \ldots, \qquad B_0 - AB_1 = a_1 I, \qquad -AB_0 = a_0 I$$

Multiplicando essas equações por $A^n, \; A^{n-1}, \; \ldots, \; A, \; I$, respectivamente, resulta

$$A^n B_{n-1} = A_n I, \qquad A^{n-1}B_{n-2} - A^n B_{n-1} = a_{n-1}A^{n-1}, \qquad \ldots, \qquad AB_0 - A^2 B_1 = a_1 A, \qquad -AB_0 = a_0 I$$

Somando essas equações matriciais, obtemos 0 do lado esquerdo e $\Delta(A)$ do lado direito, ou seja,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

Assim, $\Delta(A) = 0$, demonstrando o teorema de Cayley-Hamilton.

Autovalores e autovetores de matrizes 2 × 2

9.9 Seja $A = \begin{bmatrix} 3 & -4 \\ 2 & -6 \end{bmatrix}$.

(a) Encontre todos os autovalores e autovetores associados.

(b) Encontre matrizes P e D tais que P seja não singular e $D = P^{-1}AP$ seja diagonal.

(a) Começamos calculando o polinômio característico $\Delta(t)$ de A.

$$\Delta(t) = t^2 - \text{tr}(A)\ t + |A| = t^2 + 3t - 10 = (t-2)(t+5)$$

As raízes $\lambda = 2$ e $\lambda = -5$ de $\Delta(t)$ são os autovalores de A. Calculemos os autovetores associados.

(i) Subtraímos $\lambda = 2$ das entradas diagonais de A para obter a matriz $M = A - 2I$, cujo sistema homogêneo correspondente $MX = 0$ fornece os autovetores associados a $\lambda = 2$. Temos

$$M = \begin{bmatrix} 1 & -4 \\ 2 & -8 \end{bmatrix}, \qquad \text{que corresponde a} \qquad \begin{aligned} x - 4y &= 0 \\ 2x - 8y &= 0 \end{aligned} \qquad \text{ou} \qquad x - 4y = 0$$

O sistema só tem uma variável livre e $v_1 = (4, 1)$ é uma solução não nula. Assim, $v_1 = (4, 1)$ é um autovetor associado a (e gerando o subespaço de) $\lambda = 2$.

(ii) Subtraímos $\lambda = -5$ (ou, equivalentemente, somamos 5) das entradas diagonais de A para obter

$$M = \begin{bmatrix} 8 & -4 \\ 2 & -1 \end{bmatrix}, \qquad \text{que corresponde a} \qquad \begin{aligned} 8x - 4y &= 0 \\ 2x - \ \ y &= 0 \end{aligned} \qquad \text{ou} \qquad 2x - y = 0$$

O sistema só tem uma variável livre e $v_2 = (1, 2)$ é uma solução não nula. Assim, $v_2 = (1, 2)$ é um autovetor associado a $\lambda = 5$.

(b) Seja P a matriz de colunas v_1 e v_2. Então

$$P = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{e} \qquad D = P^{-1}AP = \begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}$$

Observe que D é a matriz diagonal cujas entradas diagonais são os autovalores de A associados aos autovetores que aparecem em P.

OBSERVAÇÃO Aqui P é a matriz de mudança de base da base canônica de $\mathbf{R}^2$ para a base $S = \{v_1, v_2\}$ e D é a matriz que representa (o operador matricial) A em relação à nova base S.

9.10 Seja $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$.

(a) Encontre todos os autovalores e autovetores associados.

(b) Encontre uma matriz não singular P tal que $D = P^{-1}AP$ seja diagonal e P^{-1}.

(c) Encontre A^6 e $f(A)$, com $t^4 - 3t^3 - 6t^2 + 7t + 3$.

(d) Encontre uma "raiz cúbica real" de A, isto é, uma matriz B tal que $B^3 = A$ e B tenha autovalores reais.

(a) Começamos calculando o polinômio característico $\Delta(t)$ de A.

$$\Delta(t) = t^2 - \text{tr}(A)\ t + |A| = t^2 - 5t + 4 = (t-1)(t-4)$$

As raízes $\lambda = 1$ e $\lambda = 4$ de $\Delta(t)$ são os autovalores de A. Calculemos os autovetores associados.

(i) Subtraímos $\lambda = 1$ das entradas diagonais de A para obter a matriz $M = A - I$, cujo sistema homogêneo correspondente $MX = 0$ fornece os autovetores associados a $\lambda = 1$. Temos

$$M = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \qquad \text{que corresponde a} \qquad \begin{aligned} x + 2y &= 0 \\ x + 2y &= 0 \end{aligned} \qquad \text{ou} \qquad x + 2y = 0$$

O sistema só tem uma solução independente, por exemplo, $x = 2$, $y = -1$. Assim, $v_1 = (2, -1)$ é um autovetor associado a (e gerando o subespaço de) $\lambda = 1$.

(ii) Subtraímos $\lambda = 4$ das entradas diagonais de A para obter

$$M = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}, \quad \text{que corresponde a} \quad \begin{aligned} -2x + 2y &= 0 \\ x - y &= 0 \end{aligned} \quad \text{ou} \quad x - y = 0$$

O sistema só tem uma solução independente, por exemplo, $x = 1$, $y = 1$. Assim, $v_2 = (1, 1)$ é um autovetor associado a $\lambda = 4$.

(b) Seja P a matriz de colunas v_1 e v_2. Então

$$P = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{e} \quad D = P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \text{com} \quad P^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

(c) Usamos a fatoração diagonal $A = PDP^{-1}$ e, como $1^6 = 1$ e $4^6 = 4096$, obtemos

$$A^6 = PD^6P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4096 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1366 & 2230 \\ 1365 & 2731 \end{bmatrix}$$

Também $f(1) = 2$ e $f(4) = -1$. Logo,

$$f(A) = Pf(D)P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$$

(d) Aqui, $\begin{bmatrix} 1 & 0 \\ 0 & \sqrt[3]{4} \end{bmatrix}$ é a raiz cúbica real de D. Logo, a raiz cúbica real de A é

$$B = P\sqrt[3]{D}P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt[3]{4} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 + \sqrt[3]{4} & -2 + 2\sqrt[3]{4} \\ -1 + \sqrt[3]{4} & 1 + 2\sqrt[3]{4} \end{bmatrix}$$

9.11 Cada uma das matrizes reais dadas define um operador linear de $\mathbf{R}^2$.

(a) $A = \begin{bmatrix} 5 & 6 \\ 3 & -2 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$, (c) $C = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$

Para cada matriz, encontre todos os autovalores e um conjunto S máximo de autovetores linearmente independentes. Qual desses operadores é diagonalizável, ou seja, pode ser representado por uma matriz diagonal?

(a) Começamos calculando $\Delta(t) = t^2 - 3t - 28 = (t - 7)(t + 4)$. As raízes $\lambda = 7$ e $\lambda = -4$ são os autovalores de A. Calculemos os autovetores associados.

(i) Subtraímos $\lambda = 7$ das entradas diagonais de A para obter

$$M = \begin{bmatrix} -2 & 6 \\ 3 & -9 \end{bmatrix}, \quad \text{que corresponde a} \quad \begin{aligned} -2x + 6y &= 0 \\ 3x - 9y &= 0 \end{aligned} \quad \text{ou} \quad x - 3y = 0$$

A única solução não nula é $v_1 = (3, 1)$.

(ii) Subtraímos $\lambda = -4$ (ou somamos 4) das entradas diagonais de A para obter

$$M = \begin{bmatrix} 9 & 6 \\ 3 & 2 \end{bmatrix}, \quad \text{que corresponde a} \quad \begin{aligned} 9x + 6y &= 0 \\ 3x + 2y &= 0 \end{aligned} \quad \text{ou} \quad 3x + 2y = 0$$

A única solução não nula é $v_2 = (2, -3)$.

Então $S = \{v_1, v_2\} = \{(3, 1), (2, -3)\}$ é um conjunto máximo de autovetores linearmente independentes. Como S é uma base de $\mathbf{R}^2$, A é diagonalizável. Em relação à base S, A é representada pela matriz diagonal $D = \text{diag}(7, -4)$.

(b) Começamos calculando o polinômio característico $\Delta(t) = t^2 + 1$. Não há raízes reais. Assim, B é uma matriz real que representa um operador linear de $\mathbf{R}^2$ que não tem autovalores e não tem autovetores. Logo, em particular, B não é diagonalizável.

(c) Começamos calculando $\Delta(t) = t^2 - 8t + 16 = (t-4)^2$. Assim, $\lambda = 4$ é o único autovalor de C. Subtraímos $\lambda = 4$ das entradas diagonais de C para obter

$$M = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \qquad \text{que corresponde a} \qquad x - y = 0$$

O sistema homogêneo tem uma única solução independente, por exemplo, $x = 1$, $y = 1$. Assim, $v = (1, 1)$ é um autovetor de C. Mais que isso, como não há outros autovalores, o conjunto $S = \{v\} = \{(1, 1)\}$ de um elemento é um conjunto máximo de autovetores linearmente independentes de C. Em particular, como S não é uma base de $\mathbf{R}^2$, C não é diagonalizável.

9.12 Considere o operador linear B do espaço complexo $\mathbf{C}^2$ definido pela matriz B do Problema 9.11. Mostre que, nesse caso, B é diagonalizável, encontrando uma base S de $\mathbf{C}^2$ consistindo em autovetores de B.

O polinômio característico de B ainda é $\Delta(t) = t^2 + 1$. Como um polinômio sobre $\mathbf{C}$, $\Delta(t)$ fatora, já que $\Delta(t) = (t-i)(t+i)$. Assim, $\lambda = i$ e $\lambda = -i$ são os autovalores de B.

(i) Subtraímos $\lambda = i$ das entradas diagonais de B para obter o sistema homogêneo

$$\begin{aligned} (1-i)x - \qquad\quad y &= 0 \\ 2x + (-1-i)y &= 0 \end{aligned} \qquad \text{ou} \qquad (1-i)x - y = 0$$

O sistema só tem uma solução independente, por exemplo, $x = 2$, $y = 1 - i$. Assim, $v_1 = (1, \; 1-i)$ é um autovetor que gera o subespaço de $\lambda = i$.

(ii) Subtraímos $\lambda = -i$ (ou somamos i) das entradas diagonais de B para obter o sistema homogêneo

$$\begin{aligned} (1+i)x - \qquad\quad y &= 0 \\ 2x + (-1+i)y &= 0 \end{aligned} \qquad \text{ou} \qquad (1+i)x - y = 0$$

O sistema só tem uma solução independente, por exemplo, $x = 1$, $y = 1 + i$. Assim, $v_2 = (1, \; 1+i)$ é um autovetor que gera o autoespaço de $\lambda = -i$.

Como uma matriz complexa, B é diagonalizável. Mais precisamente, $S = \{v_1, v_2\} = \{(1, 1 - i), (1, 1 + i)\}$ é uma base de $\mathbf{C}^2$ consistindo em autovetores de B. Usando essa base S, B é representada pela matriz diagonal $D = \text{diag}(i, -i)$.

9.13 Seja L o operador linear de $\mathbf{R}^2$ que reflete cada ponto P pela reta $y = kx$, com $k > 0$ fixado. (Ver Figura 9-1.)

(a) Mostre que $v_1 = (k, 1)$ e $v_2 = (1, -k)$ são autovetores de L.

(b) Mostre que L é diagonalizável e encontre uma representação diagonal D de L.

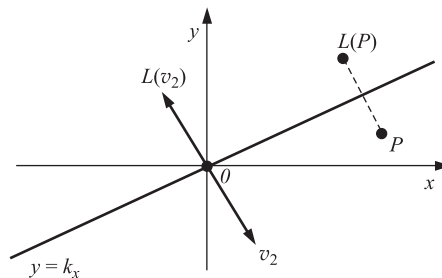



Figura 9-1

(a) O vetor $v_1 = (k, 1)$ determina um ponto da reta $y = kx$ e, portanto, é mantido fixado por L, isto é, $L(v_1) = v_1$. Assim, v_1 é um autovetor de L associado ao autovalor $\lambda_1 = 1$.

O vetor $v_2 = (1, -k)$ é perpendicular à reta $y = kx$ e, portanto, L reflete v_2 em seu simétrico, isto é, $L(v_2) = -v_2$. Assim, v_2 é um autovetor de L associado ao autovalor $\lambda_2 = -1$.

(b) Dessa forma, $S = \{v_1, v_2\}$ é uma base de $\mathbf{R}^2$ consistindo em autovetores de L. Assim, L é diagonalizável, com a representação matricial $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (em relação à base S).

Autovalores e autovetores

9.14 Seja $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$. (a) Encontre todos os autovalores de A.

(b) Encontre um conjunto S máximo de autovetores linearmente independentes de A.

(c) A é diagonalizável? Se for, encontre P tal que $D = P^{-1}AP$ seja diagonal.

(a) Começamos calculando o polinômio característico $\Delta(t)$ de A. Temos

$$\text{tr}(A) = 4 + 5 + 2 = 11 \qquad \text{e} \qquad |A| = 40 - 2 - 2 + 5 + 8 - 4 = 45$$

Também calculamos cada cofator A_{ii} de a_{ii} em A.

$$A_{11} = \begin{vmatrix} 5 & -2 \\ 1 & 2 \end{vmatrix} = 12, \qquad A_{22} = \begin{vmatrix} 4 & -1 \\ 1 & 2 \end{vmatrix} = 9, \qquad A_{33} = \begin{vmatrix} 4 & 1 \\ 2 & 5 \end{vmatrix} = 18$$

Logo,

$$\Delta(t) = t^3 - \text{tr}(A)\, t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 11t^2 + 39t - 45$$

Supondo que $\Delta(t)$ tenha alguma raiz racional, os candidatos são $\pm 1,\ \pm 3,\ \pm 5,\ \pm 9,\ \pm 15,\ \pm 45$. Testando por Ruffini, obtemos

$$\begin{array}{r|r} 3 & 1 - 11 + 39 - 45 \\ & 3 - 24 + 45 \\ \hline & 1 - 8 + 15 + \ \ 0 \end{array}$$

Assim, $t = 3$ é uma raiz de $\Delta(t)$. Também $t - 3$ e $t^2 - 8t + 15$ são fatores. Logo,

$$\Delta(t) = (t - 3)(t^2 - 8t + 15) = (t - 3)(t - 5)(t - 3) = (t - 3)^2(t - 5)$$

Por isso, $\lambda = 3$ e $\lambda = 5$ são autovalores de A.

(b) Calculamos autovetores linearmente independentes para cada autovalor de A.

(i) Subtraímos $\lambda = 3$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{bmatrix}, \qquad \text{correspondente a} \qquad x + y - z = 0$$

Logo, $u = (1, -1, 0)$ e $v = (1, 0, 1)$ são soluções linearmente independentes.

(ii) Subtraímos $\lambda = 5$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{bmatrix} \qquad \text{correspondente a} \qquad \begin{aligned} -x + y - \ \ z &= 0 \\ 2x - \qquad 2z &= 0 \\ x + y - 3z &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x \qquad - \ \ z &= 0 \\ y - 2z &= 0 \end{aligned}$$

A única variável livre é z. Logo, $w = (1, 2, 1)$ é uma solução.

Assim, $S = \{u, v, w\} = \{(1, -1, 0), (1, 0, 1), (1, 2, 1)\}$ é um conjunto máximo de autovetores de A linearmente independentes.

OBSERVAÇÃO Os vetores u e v foram escolhidos como soluções independentes do sistema dado por $x + y - z = 0$. Por outro lado, w é, automaticamente, independente de u e v, porque w pertence a um autoespaço diferente de A. Assim, os três vetores são linearmente independentes.

(c) A é diagonalizável porque tem três autovetores linearmente independentes. Seja P a matriz de colunas u, v, w. Então

$$P = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{e} \quad D = P^{-1}AP = \begin{bmatrix} 3 & & \\ & 3 & \\ & & 5 \end{bmatrix}$$

9.15 Repita o Problema 9.14 com a matriz $B = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$.

(a) Começamos calculando o polinômio característico $\Delta(t)$ de B. Temos

$$\text{tr}(B) = 0, \quad |B| = -16, \quad B_{11} = -4, \quad B_{22} = 0, \quad B_{33} = -8, \quad \text{portanto,} \quad \sum_i B_{ii} = -12$$

Logo, $\Delta(t) = t^3 - 12t + 16 = (t-2)^2(t+4)$. Assim, $\lambda_1 = 2$ e $\lambda_2 = -4$ são os autovalores de B.

(b) Calculamos uma base do autoespaço de cada autovalor de B.

(i) Subtraímos $\lambda_1 = 2$ das entradas diagonais de B para obter

$$M = \begin{bmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{bmatrix} \quad \text{correspondente a} \quad \begin{aligned} x - y + z &= 0 \\ 7x - 7y + z &= 0 \\ 6x - 6y \quad &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + z &= 0 \\ z &= 0 \end{aligned}$$

O sistema só tem uma solução independente, por exemplo, $x = 1$, $y = 1$, $z = 0$. Assim, $u = (1, 1, 0)$ constitui uma base do autoespaço de $\lambda_1 = 2$.

(ii) Subtraímos $\lambda_2 = -4$ (ou somamos 4) das entradas diagonais de B para obter

$$M = \begin{bmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{bmatrix} \quad \text{correspondente a} \quad \begin{aligned} 7x - y + z &= 0 \\ 7x - y + z &= 0 \\ 6x - 6y + 6z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + z &= 0 \\ 6y - 6z &= 0 \end{aligned}$$

O sistema só tem uma solução independente, por exemplo, $x = 0$, $y = 1$, $z = 1$. Assim, $v = (0, 1, 1)$ constitui uma base do autoespaço de $\lambda_2 = -4$.

Assim, $S = \{u, v\}$ é um conjunto máximo de autovetores de B linearmente independentes.

(c) Como B tem, no máximo, dois autovetores linearmente independentes, B não é semelhante a uma matriz diagonal, ou seja, B não é diagonalizável.

9.16 Encontre as multiplicidades algébrica e geométrica do autovalor $\lambda_1 = 2$ da matriz B do Problema 9.15.

A multiplicidade algébrica de $\lambda_1 = 2$ é 2, porque $t - 2$ aparece com expoente 2 em $\Delta(t)$. Contudo, a multiplicidade geométrica de $\lambda_1 = 2$ é 1, porque $E_{\lambda_1} = 1$ (sendo E_{λ_1} o autoespaço de λ_1).

9.17 Seja $T: \mathbf{R}^3 \to \mathbf{R}^3$ definido por $T(x, y, z) = (2x + y - 2z, \quad 2x + 3y - 4z, \quad x + y - z)$. Encontre todos os autovalores de T e uma base em cada autoespaço. T é diagonalizável? Se for, encontre uma base S de $\mathbf{R}^3$ que diagonaliza T e encontre sua representação diagonal.

Inicialmente encontramos a matriz A que representa T na base canônica de $\mathbf{R}^3$, escrevendo os coeficientes de x, y, z como linhas e depois calculamos o polinômio característico de A (e de T). Temos

$$A = [T] = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{bmatrix} \quad \text{e} \quad \begin{gathered} \text{tr}(A) = 4, \quad |A| = 2 \\ A_{11} = 1, \quad A_{22} = 0, \quad A_{33} = 4 \\ \sum_i A_{ii} = 5 \end{gathered}$$

Logo, $\Delta(t) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2)$ e, portanto, $\lambda = 1$ e $\lambda = 2$ são os autovalores de A (e de T). Agora, calculamos autovetores linearmente independentes de cada autoespaço de A.

(i) Subtraímos $\lambda = 1$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} 1 & 1 & -2 \\ 2 & 2 & -4 \\ 1 & 1 & -2 \end{bmatrix} \quad \text{correspondente a} \quad x + y - 2z = 0$$

As variáveis livres são y e z, portanto existem dois autovetores linearmente independentes associados a $\lambda = 1$. Por exemplo, $u = (1, -1, 0)$ e $v = (2, 0, 1)$ são dois desses autovetores.

(ii) Subtraímos $\lambda = 2$ das entradas diagonais de A para obter

$$M = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & -4 \\ 1 & 1 & -3 \end{bmatrix} \quad \text{correspondente a} \quad \begin{aligned} y - 2z &= 0 \\ 2x + y - 4z &= 0 \\ x + y - 3z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y - 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

A única variável livre é z. Logo, $w = (1, 2, 1)$ é uma solução.

Assim, T é diagonalizável, porque tem três autovetores linearmente independentes. Mais precisamente, escolhendo

$$S = \{u, v, w\} = \{(1, -1, 0), (2, 0, 1), (1, 2, 1)\}$$

como base, T é representado pela matriz diagonal $D = \text{diag}(1, 1, 2)$.

9.18 Mostre as afirmações seguintes relativas a um operador linear (matriz) T.

(a) O escalar 0 é um autovalor de T se, e só se, T é singular.

(b) Se λ é um autovalor de T e T é invertível, então λ^{-1} é um autovalor de T^{-1}.

(a) Temos que 0 é um autovalor de T se, e só se, existe um vetor $v \neq 0$ tal que $T(v) = 0v = 0$, isto é, se, e só se, T é singular.

(b) Sendo invertível, T é não singular. Logo, por (a), $\lambda \neq 0$. Pela definição de autovalor, existe $v \neq 0$ tal que $T(v) = \lambda v$. Aplicando T^{-1} a ambos lados, obtemos

$$v = T^{-1}(\lambda v) = \lambda T^{-1}(v) \quad \text{e, portanto,} \quad T^{-1}(v) = \lambda^{-1} v$$

Assim, λ^{-1} é um autovalor de T^{-1}.

9.19 Seja λ um autovalor de um operador linear $T: V \to V$ e seja E_λ o conjunto de todos autovetores associados a λ (denominado *autoespaço* de λ). Prove que E_λ é um subespaço de V, ou seja, prove as afirmações seguintes.

(a) Se $u \in E_\lambda$, então $ru \in E_\lambda$, para cada escalar r. (b) Se $u, v, \in E_\lambda$, então $u + v \in E_\lambda$.

(a) Como $u \in E_\lambda$, temos $T(u) = \lambda u$. Então $T(ru) = rT(u) = r(\lambda u) = \lambda(ru)$, de modo que $ru \in E_\lambda$.

(Interpretamos o vetor nulo $0 \in V$ como um "autovetor" de λ para que E_λ seja um subespaço de V.)

(b) Como $u, v \in E_\lambda$, temos $T(u) = \lambda u$ e $T(v) = \lambda v$. Então $T(u + v) = T(u) + T(v) = \lambda u + \lambda v = \lambda(u + v)$ e, portanto, $u + v \in E_\lambda$.

9.20 Demonstre o Teorema 9.6. As afirmações dadas são equivalentes. (i) O escalar λ é um autovalor de A.

(ii) A matriz $\lambda I - A$ é singular.

(iii) O escalar λ é uma raiz do polinômio característico $\Delta(t)$ de A.

O escalar λ é um autovalor de A se, e só se, existe um vetor não nulo v tal que

$$Av = \lambda v \quad \text{ou} \quad (\lambda I)v - Av = 0 \quad \text{ou} \quad (\lambda I - A)v = 0$$

ou $\lambda I - A$ é singular. Nesse caso, λ é uma raiz de $\Delta(t) = |tI - A|$. Também, v está no subespaço E_λ de λ se, e só se, valer essa relação. Logo, v é uma solução de $(\lambda I - A)X = 0$.

9.21 Demonstre o Teorema 9.8′. Suponha que $v_1, v_2, ..., v_n$ sejam autovetores de T associados a autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Então $v_1, v_2, ..., v_n$ são linearmente independentes.

Vamos supor que o teorema seja falso. Seja $v_1, v_2, ..., v_s$ um conjunto mínimo de vetores para o qual o teorema seja falso. Temos $s > 1$, porque $v_1 \neq 0$, já que é autovetor. Também, pela condição de minimalidade, $v_2, ..., v_s$ é linearmente independente. Assim, v_1 é uma combinação linear de $v_2, ..., v_s$, digamos,

$$v_1 = a_2 v_2 + a_3 v_3 + \cdots + a_s v_s \qquad (1)$$

(onde algum $a_k \neq 0$). Aplicando T a (1) e usando a linearidade de T, obtemos

$$T(v_1) = T(a_2 v_2 + a_3 v_3 + \cdots + a_s v_s) = a_2 T(v_2) + a_3 T(v_3) + \cdots + a_s T(v_s) \qquad (2)$$

Como v_j é um autovetor de T associado a λ_j, temos $T(v_j) = \lambda_j v_j$. Substituindo em (2), resulta

$$\lambda_1 v_1 = a_2 \lambda_2 v_2 + a_3 \lambda_3 v_3 + \cdots + a_s \lambda_s v_s \qquad (3)$$

Multiplicando (1) por λ_1, obtemos

$$\lambda_1 v_1 = a_2 \lambda_1 v_2 + a_3 \lambda_1 v_3 + \cdots + a_s \lambda_1 v_s \qquad (4)$$

Igualando os lados direitos de (3) e (4), ou subtraindo (3) de (4), resulta

$$a_2(\lambda_1 - \lambda_2) v_2 + a_3(\lambda_1 - \lambda_3) v_3 + \cdots + a_s(\lambda_1 - \lambda_s) v_s = 0 \qquad (5)$$

Como $v_2, v_3, ..., v_s$ são linearmente independentes, os coeficientes de (5) devem ser, todos, nulos, ou seja,

$$a_2(\lambda_1 - \lambda_2) = 0, \qquad a_3(\lambda_1 - \lambda_3) = 0, \qquad \ldots, \qquad a_s(\lambda_1 - \lambda_s) = 0$$

Contudo, os λ_i são distintos. Logo, $\lambda_1 - \lambda_j \neq 0$ para $j > 1$. Assim, $a_2 = 0,\ a_3 = 0, \ldots, a_s = 0$. Isso contradiz o fato de que algum $a_k \neq 0$. O teorema está provado.

9.22 Demonstre o Teorema 9.9. Suponha que o polinômio característico $\Delta(t) = (t - a_1)(t - a_2)\ldots(t - a_n)$ de uma matriz quadrada A de ordem n seja um produto de n fatores distintos, digamos, a_i. Então A é semelhante à matriz diagonal $D = \text{diag}(a_1, a_2, ..., a_n)$.

Sejam $v_1, v_2, ..., v_n$ autovetores (não nulos) associados aos autovalores a_i. Então os n autovetores v_i são linearmente independentes (Teorema 9.8) e, portanto, formam uma base de K^n. Dessa forma, A é diagonalizável, isto é, A é semelhante a uma matriz diagonal D, e os elementos diagonais de D são os autovalores a_i.

9.23 Demonstre o Teorema 9.10′. A multiplicidade geométrica de um autovalor λ de T não é maior do que a multiplicidade algébrica.

Suponha que a multiplicidade geométrica de λ seja r. Então o autoespaço E_λ desse autovalor contém r autovetores linearmente independentes $v_1, ..., v_r$. Estendemos o conjunto $\{v_i\}$ a uma base de V, digamos, $\{v_i, \ldots, v_r, w_1, \ldots, w_s\}$. Então

$$T(v_1) = \lambda v_1, \qquad T(v_2) = \lambda v_2, \qquad \ldots, \qquad T(v_r) = \lambda v_r,$$

$$\begin{aligned}
T(w_1) &= a_{11} v_1 + \cdots + a_{1r} v_r + b_{11} w_1 + \cdots + b_{1s} w_s \\
T(w_2) &= a_{21} v_1 + \cdots + a_{2r} v_r + b_{21} w_1 + \cdots + b_{2s} w_s \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(w_s) &= a_{s1} v_1 + \cdots + a_{sr} v_r + b_{s1} w_1 + \cdots + b_{ss} w_s
\end{aligned}$$

A matriz de T nessa base é $M = \begin{bmatrix} \lambda I_r & A \\ 0 & B \end{bmatrix}$, com $A = [a_{ij}]^T$ e $B = [b_{ij}]^T$.

Como M é diagonal em blocos, o polinômio característico $(t - \lambda)^r$ do bloco λI_r deve dividir o polinômio característico de M e, portanto, de T. Assim, a multiplicidade algébrica de λ para T é r, no mínimo, como queríamos mostrar.

Diagonalização de matrizes reais simétricas e formas quadráticas

9.24 Seja $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$. Encontre uma matriz ortogonal P tal que $D = P^{-1}AP$ seja diagonal.

Começamos calculando o polinômio característico $\Delta(t)$ de A. Temos

$$\Delta(t) = t^2 - \text{tr}(A)\ t + |A| = t^2 - 6t - 16 = (t-8)(t+2)$$

Assim, os autovalores de A são $\lambda = 8$ e $\lambda = -2$. Em seguida, calculamos os autovetores associados. Subtraímos $\lambda = 8$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \quad \text{correspondente a} \quad \begin{matrix} -x + 3y = 0 \\ 3x - 9y = 0 \end{matrix} \quad \text{ou} \quad x - 3y = 0$$

Uma solução não nula é $u_1 = (3, 1)$.

Subtraímos $\lambda = -2$ (ou somamos 2) das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \quad \text{correspondente a} \quad \begin{matrix} 9x + 3y = 0 \\ 3x + \ \ y = 0 \end{matrix} \quad \text{ou} \quad 3x + y = 0$$

Uma solução não nula é $u_2 = (1, -3)$.
Como era de se esperar, dada a simetria de A, os autovetores u_1 e u_2 são ortogonais. Normalizamos u_1 e u_2 para obter, respectivamente, os vetores unitários

$$\hat{u}_1 = (3/\sqrt{10}, 1/\sqrt{10}) \quad \text{e} \quad \hat{u}_2 = (1/\sqrt{10}, -3/\sqrt{10}).$$

Finalmente, seja P a matriz cujas colunas são os vetores unitários $\hat{u}_1$ e $\hat{u}_2$, respectivamente. Então

$$P = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \quad \text{e} \quad D = P^{-1}AP = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}$$

Conforme previsto, as entradas diagonais de D são os autovalores de A.

9.25 Seja $B = \begin{bmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{bmatrix}$. (a) Encontre todos os autovalores de B.

(b) Encontre um conjunto S máximo de autovetores ortogonais de B.

(c) Encontre uma matriz ortogonal P tal que $D = P^{-1}BP$ seja diagonal.

(a) Começamos calculando o polinômio característico de B. Temos

$$\text{tr}(B) = 6, \quad |B| = 400, \quad B_{11} = 0, \quad B_{22} = -60, \quad B_{33} = -75, \quad \text{portanto,} \quad \sum_i B_{ii} = -135$$

Logo, $\Delta(t) = t^3 - 6t^2 - 135t - 400$. Se $\Delta(t)$ tiver alguma raiz inteira, os candidatos devem dividir 400. Testando $t = –5$, por Ruffini obtemos

$$\begin{array}{r|rrrr} -5 & 1 - & 6 - & 135 - & 400 \\ & - & 5 + & 55 + & 400 \\ \hline & 1 - & 11 - & 80 + & 0 \end{array}$$

Assim, $t + 5$ é um fator de $\Delta(t)$, e $t^2 - 11t - 80$ também é um fator. Assim,

$$\Delta(t) = (t+5)(t^2 - 11t - 80) = (t+5)^2(t-16)$$

Os autovalores de B são $\lambda = -5$ (multiplicidade 2) e $\lambda = 16$ (multiplicidade 1).

(b) Encontremos uma base ortogonal de cada autoespaço. Subtraímos $\lambda = -5$ (ou somamos 5) das entradas diagonais de B para obter o sistema homogêneo

$$16x - 8y + 4z = 0, \quad -8x + 4y - 2z = 0, \quad 4x - 2y + z = 0$$

Logo, $4x - 2y + z = 0$. O sistema tem duas soluções independentes. Uma solução é $v_1 = (0, 1, 2)$. Procuramos uma segunda solução $v_2 = (a, b, c)$ que seja ortogonal a v_1, ou seja,

$$4a - 2b + c = 0 \quad \text{e, portanto,} \quad b - 2c = 0$$

Uma tal solução é $v_2 = (-5, -8, 4)$.

Subtraímos $\lambda = 16$ das entradas diagonais de B para obter o sistema homogêneo

$$-5x - 8y + 4z = 0, \qquad -8x - 17y - 2z = 0, \qquad 4x - 2y - 20z = 0$$

Esse sistema fornece uma solução não nula $v_3 = (4, -2, 1)$. (Como afirma o Teorema 9.13, o autovetor v_3 é ortogonal a v_1 e v_2.)

Então v_1, v_2, v_3 formam um conjunto máximo de autovetores ortogonais (não nulos) de B.

(c) Normalizamos v_1, v_2, v_3 para obter uma base ortogonal

$$\hat{v}_1 = v_1/\sqrt{5}, \qquad \hat{v}_2 = v_2/\sqrt{105}, \qquad \hat{v}_3 = v_3/\sqrt{21}$$

Então P é a matriz cujas colunas são $\hat{v}_1, \hat{v}_2, \hat{v}_3$. Assim,

$$P = \begin{bmatrix} 0 & -5/\sqrt{105} & 4/\sqrt{21} \\ 1/\sqrt{5} & -8/\sqrt{105} & -2/\sqrt{21} \\ 2/\sqrt{5} & 4/\sqrt{105} & 1/\sqrt{21} \end{bmatrix} \qquad \text{e} \qquad D = P^{-1}BP = \begin{bmatrix} -5 & & \\ & -5 & \\ & & 16 \end{bmatrix}$$

9.26 Seja $q(x, y) = x^2 + 6xy - 7y^2$. Encontre uma substituição ortogonal que diagonalize q.

Começamos com a matriz simétrica A que representa q e seu polinômio característico $\Delta(t)$. Temos

$$A = \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \qquad \text{e} \qquad \Delta(t) = t^2 + 6t - 16 = (t-2)(t+8)$$

Os autovalores de A são $\lambda = 2$ e $\lambda = -8$. Assim, usando as novas variáveis s e t, uma forma diagonal de q é

$$q(s, t) = 2s^2 - 8t^2$$

A substituição ortogonal correspondente é obtida a partir de um conjunto ortogonal de autovetores de A.

(i) Subtraímos $\lambda = 2$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \text{correspondente a} \begin{matrix} -x + 3y = 0 \\ 3x - 9y = 0 \end{matrix} \qquad \text{ou} \qquad -x + 3y = 0$$

Uma solução não nula é $u_1 = (3, 1)$.

(ii) Subtraímos $\lambda = -8$ (ou somamos 8) das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \qquad \text{correspondente a} \qquad \begin{matrix} 9x + 3y = 0 \\ 3x + \ \ y = 0 \end{matrix} \qquad \text{ou} \qquad 3x + y = 0$$

Uma solução não nula é $u_2 = (-1, 3)$.

Como era de se esperar, dada a simetria de A, os autovetores u_1 e u_2 são ortogonais.

Agora normalizamos u_1 e u_2 para obter, respectivamente, os vetores unitários

$$\hat{u}_1 = (3/\sqrt{10},\ 1/\sqrt{10}) \qquad \text{e} \qquad \hat{u}_2 = (-1/\sqrt{10},\ 3/\sqrt{10}).$$

Finalmente, seja P a matriz cujas colunas são os vetores unitários $\hat{u}_1$ e $\hat{u}_2$, respectivamente. Então $[x, y]^T = P[s, t]^T$ é a mudança de coordenadas ortogonal procurada, ou seja,

$$P = \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix} \qquad \text{e} \qquad x = \frac{3s - t}{\sqrt{10}}, \qquad y = \frac{s + 3t}{\sqrt{10}}$$

Também podemos expressar s e t em termos de x e y usando $P^{-1} = P^T$, ou seja,

$$s = \frac{3x + y}{\sqrt{10}}, \qquad t = \frac{-x + 3t}{\sqrt{10}}$$

Polinômio mínimo

9.27 Sejam $A = \begin{bmatrix} 4 & -2 & 2 \\ 6 & -3 & 4 \\ 3 & -2 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & -2 & 2 \\ 4 & -4 & 6 \\ 2 & -3 & 5 \end{bmatrix}$. O polinômio característico de ambas matrizes é $\Delta(t) = (t-2)(t-1)^2$. Encontre o polinômio mínimo $m(t)$ de cada matriz.

O polinômio mínimo $m(t)$ deve dividir $\Delta(t)$. Também, cada fator de $\Delta(t)$ (ou seja, $t-2$ e $t-1$) deve ser, também, fator de $m(t)$. Assim, $m(t)$ é exatamente um dos dois polinômios seguintes.

$$f(t) = (t-2)(t-1) \qquad \text{ou} \qquad g(t) = (t-2)(t-1)^2$$

(a) Pelo Teorema de Cayley-Hamilton, $g(A) = \Delta(A) = 0$, de modo que basta testar $f(t)$. Temos

$$f(A) = (A-2I)(A-I) = \begin{bmatrix} 2 & -2 & 2 \\ 6 & -5 & 4 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 & 2 \\ 6 & -4 & 4 \\ 3 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, $m(t) = f(t) = (t-2)(t-1) = t^2 - 3t + 2$ é o polinômio mínimo de A.

(b) Novamente, $g(B) = \Delta(B) = 0$, de modo que basta testar $f(t)$. Obtemos

$$f(B) = (B-2I)(B-I) = \begin{bmatrix} 1 & -2 & 2 \\ 4 & -6 & 6 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} 2 & -2 & 2 \\ 4 & -5 & 6 \\ 2 & -3 & 4 \end{bmatrix} = \begin{bmatrix} -2 & 2 & -2 \\ -4 & 4 & -4 \\ -2 & 2 & -2 \end{bmatrix} \neq 0$$

Assim, $m(t) \neq f(t)$. Em vista disso, $m(t) = g(t) = (t-2)(t-1)^2$ é o polinômio mínimo de B. [Enfatizamos que não é necessário calcular $g(B)$, pois o Teorema de Cayley-Hamilton garante que $g(B) = 0$.]

9.28 Encontre o polinômio mínimo $m(t)$ de cada uma das matrizes dadas.

(a) $A = \begin{bmatrix} 5 & 1 \\ 3 & 7 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$, (c) $C = \begin{bmatrix} 4 & -1 \\ 1 & 2 \end{bmatrix}$

(a) O polinômio característico de A é $\Delta(t) = t^2 - 12t + 32 = (t-4)(t-8)$. Como $\Delta(t)$ tem fatores distintos, o polinômio mínimo $m(t) = \Delta(t) = t^2 - 12t + 32$.

(b) Como B é triangular, seus autovalores são os elementos diagonais 1, 2, 3 e, então, seu polinômio característico é $\Delta(t) = (t-1)(t-2)(t-3)$. Como $\Delta(t)$ tem fatores distintos, $m(t) = \Delta(t)$.

(c) O polinômio característico de C é $\Delta(t) = t^2 - 6t + 9 = (t-3)^2$. Logo, o polinômio mínimo de C é $f(t) = t - 3$ ou $g(t) = (t-3)^2$. Contudo, $f(C) \neq 0$, ou seja, $C - 3I \neq 0$. Logo,

$$m(t) = g(t) = \Delta(t) = (t-3)^2.$$

9.29 Seja $S = \{u_1, u_2, \ldots, u_n\}$ uma base de V e suponha que F e G sejam operadores de V tais que $[F]$ tem zeros na diagonal e abaixo dela e que $[G]$ tem $a \neq 0$ na superdiagonal e zeros no resto. Ou seja,

$$[F] = \begin{bmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \qquad [G] = \begin{bmatrix} 0 & a & 0 & \ldots & 0 \\ 0 & 0 & a & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

Mostre que (a) $F^n = 0$, (b) $G^{n-1} \neq 0$, mas $G^n = 0$. (Essas condições também valem para $[F]$ e $[G]$.)

(a) Temos $F(u_1) = 0$ e, para $r > 1$, $F(u_r)$ é uma combinação linear de vetores que precedem u_r em S. Ou seja,

$$F(u_r) = a_{r1}u_1 + a_{r2}u_2 + \cdots + a_{r,r-1}u_{r-1}$$

Então, $F^2(u_r) = F(F(u_r))$ é uma combinação linear de vetores que precedem u_{r-1} e assim por diante. Logo, $F^r(u_r) = 0$ para cada r. Assim, para cada r, $F^n(u_r) = F^{n-r}(0) = 0$ e, portanto, $F^n = 0$, como queríamos mostrar.

(b) Temos $G(u_1) = 0$ e, para cada $k > 1$, $G(u_k) = au_{k-1}$. Logo, $G^r(u_k) = a^r u_{k-r}$ para $r < k$. Como $a \neq 0$, também $a^{n-1} \neq 0$. Portanto, $G^{n-1}(u_n) = a^{n-1}u_1 \neq 0$ e, assim, $G^{n-1} \neq 0$. Por outro lado, (a) afirma que $G^n = 0$.

9.30 Seja A a matriz do Exemplo 9.12(a) que tem um λ em cada entrada diagonal, a em cada entrada superdiagonal, com $a \neq 0$, e zeros nas demais. Mostre que $f(t) = (t - \lambda)^n$ é tanto o polinômio característico quanto o polinômio mínimo $m(t)$ de A.

Como A é triangular, com λ na diagonal, $\Delta(t) = f(t) = (t - \lambda)^n$ é o polinômio característico de A. Assim, $m(t)$ é uma potência de $t - \lambda$. Pelo Problema 9.29, $(A - \lambda I)^{r-1} \neq 0$. Logo, $m(t) = \Delta(t) = (t - \lambda)^n$.

9.31 Encontre o polinômio característico $\Delta(t)$ e o polinômio mínimo $m(t)$ de cada matriz dada.

(a) $M = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$, (b) $M' = \begin{bmatrix} 2 & 7 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{bmatrix}$

(a) M é diagonal em blocos com blocos diagonais

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

O polinômio característico e mínimo de A é $f(t) = (t - 4)^3$ e o polinômio característico e mínimo de B é $g(t) = (t - 4)^2$. Então

$$\Delta(t) = f(t)g(t) = (t - 4)^5, \quad \text{mas} \quad m(t) = \text{MMC}[f(t), g(t)] = (t - 4)^3$$

(onde MMC significa o mínimo múltiplo comum). Enfatizamos que o expoente de $m(t)$ é o tamanho do maior bloco.

(b) Aqui, M' é diagonal em blocos com blocos diagonais $A' = \begin{bmatrix} 2 & 7 \\ 0 & 2 \end{bmatrix}$ e $B' = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$. O polinômio característico de A' é $f(t) = (t - 2)^2$, que tem fatores distintos. Logo, $g(t)$ também é o polinômio mínimo de B. Em vista disso,

$$\Delta(t) = f(t)g(t) = (t - 2)^3(t - 3), \quad \text{mas} \quad m(t) = \text{MMC}[f(t), g(t)] = (t - 2)^2(t - 3)$$

9.32 Encontre uma matriz A cujo polinômio mínimo seja $f(t) = t^3 - 8t^2 + 5t + 7$.

Simplesmente tomamos $A = \begin{bmatrix} 0 & 0 & -7 \\ 1 & 0 & -5 \\ 0 & 1 & 8 \end{bmatrix}$, a matriz companheira de $f(t)$ [definida no Exemplo 9.12(b).]

9.33 Demonstre o Teorema 9.15. O polinômio mínimo $m(t)$ de uma matriz (operador linear) A divide qualquer polinômio que tenha A como raiz. Em particular (pelo Teorema de Cayley-Hamilton), $m(t)$ divide o polinômio característico $\Delta(t)$ de A.

Seja $f(t)$ um polinômio tal que $f(A) = 0$. Pelo algoritmo da divisão, existem polinômios $q(t)$ e $r(t)$ tais que $f(t) = m(t)q(t) + r(t)$ e $r(t) = 0$ ou, então, o grau de $r(t)$ é menor do que o grau de $m(t)$. Substituindo $t = A$ nessa equação e usando que $f(A) = 0$ e $m(A) = 0$, obtemos $r(A) = 0$. Se $r(t) \neq 0$, então $r(t)$ é um polinômio de grau menor do que $m(t)$ que tem A como zero. Isso contradiz a definição de polinômio mínimo. Assim, $r(t) = 0$ e, portanto, $f(t) = m(t)q(t)$, ou seja, $m(t)$ divide $f(t)$.

9.34 Seja $m(t)$ o polinômio mínimo $m(t)$ de uma matriz quadrada A de ordem n. Prove que o polinômio característico $\Delta(t)$ de A divide $[m(t)]^n$.

Suponha que $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Definimos matrizes B_j como segue.

$$\begin{array}{lll} B_0 = I & \text{logo} & I = B_0 \\ B_1 = A + c_1 I & \text{logo} & c_1 I = B_1 - A = B_1 - AB_0 \\ B_2 = A^2 + c_1 A + c_2 I & \text{logo} & c_2 I = B_2 - A(A + c_1 I) = B_2 - AB_1 \\ \cdots & & \\ B_{r-1} = A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I & \text{logo} & c_{r-1} I = B_{r-1} - AB_{r-2} \end{array}$$

Então

$$-AB_{r-1} = c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) = c_r I - m(A) = c_r I$$

Definindo

$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

obtemos

$$\begin{aligned} (tI - A)B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\ &= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} tI + c_r I = m(t)I \end{aligned}$$

Tomando o determinante de ambos lados, resulta $|tI - A||B(t)| = |m(t)I| = [m(t)]^n$. Como $|B(t)|$ é um polinômio, $|tI - A|$ divide $[m(t)]^n$, ou seja, o polinômio característico de A divide $[m(t)]^n$.

9.35 Demonstre o Teorema 9.16. O polinômio característico $\Delta(t)$ e o polinômio mínimo $m(t)$ de uma matriz A têm os mesmos fatores irredutíveis.

Seja $f(t)$ um polinômio irredutível. Se $f(t)$ divide $m(t)$, então $f(t)$ também divide $\Delta(t)$ [porque $m(t)$ divide $\Delta(t)$]. Por outro lado, se $f(t)$ divide $\Delta(t)$, então, pelo Problema 9.34, $f(t)$ também divide $[m(t)]^n$. No entanto, $f(t)$ é irredutível, portanto, $f(t)$ também divide $m(t)$. Assim, $m(t)$ e $\Delta(t)$ têm os mesmos fatores irredutíveis.

9.36 Demonstre o Teorema 9.19. O polinômio mínimo $m(t)$ de uma matriz diagonal em blocos M com blocos diagonais A_i é igual ao mínimo múltiplo comum (MMC) dos polinômios mínimos dos blocos diagonais A_i.

Demonstramos o teorema no caso $r = 2$, sendo que o caso geral decorre por indução. Seja $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, com A e B matrizes quadradas. Queremos mostrar que o polinômio mínimo $m(t)$ de M é o MMC dos polinômios mínimos $g(t)$ e $h(t)$ de A e B, respectivamente.

Como $m(t)$ é o polinômio mínimo de M, temos $M, m(M) = \begin{bmatrix} m(A) & 0 \\ 0 & m(B) \end{bmatrix} = 0$ e $m(A) = 0$ e $m(B) = 0$. Como $g(t)$ é o polinômio mínimo de A, $g(t)$ divide $m(t)$. Analogamente, $h(t)$ divide $m(t)$. Assim, $m(t)$ é um múltiplo de $g(t)$ e $h(t)$.

Seja, agora, $f(t)$ um múltiplo qualquer de $g(t)$ e $h(t)$. Então $f(M) = \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$. No entanto, $m(t)$ é o polinômio mínimo de M, de modo que $m(t)$ divide $f(t)$. Assim, $m(t)$ é o MMC de $g(t)$ e $h(t)$.

9.37 Seja $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$ o polinômio mínimo de uma matriz quadrada A de ordem n. Prove as afirmações seguintes.

(a) A é não singular se, e só se, o termo constante $a_0 \neq 0$.

(b) Se A é não singular, então A^{-1} é um polinômio em A de grau $r - 1 < n$.

(a) São equivalentes. (i) A é não singular, (ii) 0 não é uma raiz de $m(t)$, (iii) $a_0 \neq 0$. Assim, a afirmação é verdadeira.

(b) Como A é não singular, temos $a_0 \neq 0$ por (a). Temos

$$m(A) = A^r + a_{r-1}A^{r-1} + \cdots + a_1A + a_0I = 0$$

Assim, $$-\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1I)A = I$$

Dessa forma, $$A^{-1} = -\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1I)$$

Problemas Complementares

Polinômios de matrizes

9.38 Sejam $A = \begin{bmatrix} 2 & -3 \\ 5 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Encontre $f(A)$, $g(A)$, $f(B)$, $g(B)$, com $f(t) = 2t^2 - 5t + 6$ e $g(t) = t^3 - 2t^2 + t + 3$.

9.39 Seja $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Encontre A^2, A^3, A^n, com $n > 3$, e A^{-1}.

9.40 Seja $B = \begin{bmatrix} 8 & 12 & 0 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{bmatrix}$. Encontre uma matriz real A tal que $B = A^3$.

9.41 Para cada matriz dada, encontre um polinômio que tem a matriz como raiz.

(a) $A = \begin{bmatrix} 2 & 5 \\ 1 & -3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & -3 \\ 7 & -4 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$

9.42 Sejam A uma matriz quadrada e $f(t)$ um polinômio quaisquer. Prove que (a) $(P^{-1}AP)^n = P^{-1}A^nP$. (b) $f(P^{-1}AP) = P^{-1}f(A)P$. (c) $f(A^T) = [f(A)]^T$. (d) Se A é simétrica, então $f(A)$ é simétrica.

9.43 Sejam $M = \text{diag}[A_1, ..., A_r]$ uma matriz diagonal em blocos e $f(t)$ um polinômio quaisquer. Mostre que $f(M)$ é diagonal em blocos e que $f(M) = \text{diag}[f(A_1), ..., f(A_r)]$.

9.44 Sejam M uma matriz triangular em blocos com blocos diagonais $A_1, ..., A_r$ e $f(t)$ um polinômio quaisquer. Mostre que $f(M)$ é triangular em blocos com blocos diagonais $f(A_1), ..., f(A_r)$.

Autovalores e autovetores

9.45 Para cada uma das matrizes dadas, encontre todos os autovalores e autovetores linearmente independentes associados.

(a) $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & 4 \\ -1 & 6 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & -4 \\ 3 & -7 \end{bmatrix}$

Quando possível, encontre a matriz não singular P que diagonaliza a matriz.

9.46 Seja $A = \begin{bmatrix} 2 & -1 \\ -2 & 3 \end{bmatrix}$.

(a) Encontre os autovalores e autovetores associados.

(b) Encontre uma matriz não singular P tal que $D = P^{-1}AP$ seja diagonal.

(c) Encontre A^8 e $f(A)$, com $f(t) = t^4 - 5t^3 + 7t^2 - 2t + 5$.

(d) Encontre uma matriz B tal que $B^2 = A$.

9.47 Repita o Problema 9.46 com $A = \begin{bmatrix} 5 & 6 \\ -2 & -2 \end{bmatrix}$.

9.48 Para cada uma das matrizes dadas, encontre todos os autovalores e um conjunto S máximo de autovetores linearmente independentes.

(a) $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$, (b) $B = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{bmatrix}$

Quais matrizes podem ser diagonalizadas, e por quê?

9.49 Para cada um dos operadores lineares $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ dados, encontre todos os autovalores e uma base de cada autoespaço.

(a) $T(x, y) = (3x + 3y,\ x + 5y)$, (b) $T(x, y) = (3x - 13y,\ x - 3y)$.

9.50 Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ uma matriz real. Encontre condições relativas a a, b, c, d que sejam necessárias e suficientes para que A seja diagonalizável, isto é, para que A tenha dois autovetores (reais) linearmente independentes.

9.51 Mostre que as matrizes A e A^T têm os mesmos autovalores. Dê um exemplo de uma matriz 2×2 A tal que A e A^T tenham autovetores diferentes.

9.52 Seja v um autovetor dos operadores lineares F e G. Mostre que v também é um autovetor do operador linear $kF + k'G$, com k e k' escalares.

9.53 Seja v um autovetor de um operador linear T associado ao autovalor λ. Demonstre.

(a) Para $n > 0$, v é um autovetor de T^n associado a λ^n.

(b) $f(\lambda)$ é um autovetor de $f(T)$, para qualquer polinômio $f(t)$.

9.54 Seja $\lambda \neq 0$ um autovalor da composta $F \circ G$ dos operadores lineares F e G. Mostre que λ também é um autovetor da composta $G \circ F$. [*Sugestão*: mostre que $G(v)$ é um autovetor de $G \circ F$.]

9.55 Seja $E\colon V \to V$ uma projeção linear, isto é, tal que $E^2 = E$. Mostre que E é diagonalizável e que, de fato, pode ser representada pela matriz diagonal $M = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, em que r é o posto de E.

Diagonalização de matrizes reais simétricas e formas quadráticas

9.56 Para cada uma das matrizes simétricas A dadas, encontre uma matriz ortogonal P e uma matriz diagonal D tais que $D = P^{-1}AP$.

(a) $A = \begin{bmatrix} 5 & 4 \\ 4 & -1 \end{bmatrix}$, (b) $A = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}$, (c) $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$

9.57 Para cada uma das matrizes simétricas B dadas, encontre seus autovalores, um conjunto ortogonal máximo S de autovetores e uma matriz ortogonal P tal que $D = P^{-1}BP$ seja diagonal.

(a) $B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & 2 & 4 \\ 2 & 5 & 8 \\ 4 & 8 & 17 \end{bmatrix}$

9.58 Usando variáveis s e t, encontre uma substituição ortogonal que diagonalize cada uma das formas quadráticas dadas.

(a) $q(x, y) = 4x^2 + 8xy - 11y^2$, (b) $q(x, y) = 2x^2 - 6xy + 10y^2$

9.59 Para cada uma das formas quadráticas $q(x, y, z)$ dadas, encontre uma substituição ortogonal que expresse x, y, z em termos das variáveis r, s, t e encontre $q(r, s, t)$.

(a) $q(x, y, z) = 5x^2 + 3y^2 + 12xz$, (b) $q(x, y, z) = 3x^2 - 4xy + 6y^2 + 2xz - 4yz + 3z^2$

9.60 Encontre uma matriz simétrica 2×2 A com autovalores

(a) $\lambda = 1$ e $\lambda = 4$ e autovetor $u = (1, 1)$ associado a $\lambda = 1$;

(b) $\lambda = 2$ e $\lambda = 3$ e autovetor $u = (1, 2)$ associado a $\lambda = 2$.

Em cada caso, encontre uma matriz B tal que $B^2 = A$.

Polinômios característico e mínimo

9.61 Encontre os polinômios característico e mínimo de cada uma das matrizes dadas.

(a) $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 4 & -2 \\ -1 & -1 & 3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 3 & 2 & -1 \\ 3 & 8 & -3 \\ 3 & 6 & -1 \end{bmatrix}$

9.62 Encontre os polinômios característico e mínimo de cada uma das matrizes dadas.

(a) $A = \begin{bmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}$, (b) $B = \begin{bmatrix} 4 & -1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$, (c) $C = \begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

9.63 Sejam $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Mostre que A e B têm polinômios característicos diferentes (portanto, não são semelhantes) mas têm o mesmo polinômio mínimo. Assim, matrizes não semelhantes podem ter o mesmo polinômio mínimo.

9.64 Seja A uma matriz quadrada de ordem n tal que $A^k = 0$, para algum $k > n$. Mostre que $A^n = 0$.

9.65 Mostre que uma matriz A e sua transposta A^T têm o mesmo polinômio mínimo.

9.66 Suponha que $f(t)$ seja um polinômio mônico irredutível tal que $f(A) = 0$ para alguma matriz A. Mostre que $f(t)$ é o polinômio mínimo de A.

9.67 Mostre que A é uma matriz escalar kI se, e só se, o polinômio mínimo de A é $m(t) = t - k$.

9.68 Encontre uma matriz A cujo polinômio mínimo seja (a) $t^3 - 5t^2 + 6t + 8$, (b) $t^4 - 5t^3 - 2t + 7t + 4$.

9.69 Sejam $f(t)$ e $g(t)$ polinômios mônicos (coeficiente dominante 1) de grau mínimo para os quais A é uma raiz. Mostre que $f(t) = g(t)$. [Assim, o polinômio mínimo de A é único.]

Respostas dos Problemas Complementares

Notação: $M = [R_1; \quad R_2; \quad \ldots]$ denota uma matriz de linhas R_1, R_2, ...

9.38 $f(A) = [-26, -3; \quad 5, -27]$, $g(A) = [-40, 39; \quad -65, -27]$,
$f(B) = [3, 6; \quad 0, 9]$, $g(B) = [3, 12; \quad 0, 15]$

9.39 $A^2 = [1, 4; \quad 0, 1]$, $A^3 = [1, 6; \quad 0, 1]$, $A^n = [1, 2n; \quad 0, 1]$, $A^{-1} = [1, -2; \quad 0, 1]$

9.40 Seja $A = [2, a, b; \quad 0, 2, c; \quad 0, 0, 2]$. Tomamos $B = A^3$ e, então, $a = 1$, $b = -\frac{1}{2}$, $c = 1$.

9.41 (a) t^2+t-11, (b) $t^2+2t+13$, (c) t^3-7t^2+6t-1

9.45 (a) $\lambda=1, u=(3,1)$; $\lambda=-4, v=(1,2)$, (b) $\lambda=4, u=(2,1)$,
(c) $\lambda=-1, u=(2,1)$; $\lambda=-5, v=(2,3)$. Somente A e C podem ser diagonalizadas; usamos $P=[u,v]$.

9.46 (a) $\lambda=1, u=(1,1)$; $\lambda=4, v=(1,-2)$,
(b) $P=[u,v]$,
(c) $f(A)=[3,1;\quad 2,1]$, $A^8=[21\ 846, -21\ 845;\quad -43\ 690, 43\ 691]$,
(d) $B=\left[\frac{4}{3},-\frac{1}{3};\quad -\frac{2}{3},\frac{5}{3}\right]$

9.47 (a) $\lambda=1, u=(3,-2)$; $\lambda=2, v=(2,-1)$, (b) $P=[u,v]$,
(c) $f(A)=[2,-6;\quad 2,9]$, $A^8=[1021, 1530;\quad -510,-764]$,
(d) $B=[-3+4\sqrt{2},\quad -6+6\sqrt{2};\quad 2-2\sqrt{2},\quad 4-3\sqrt{2}]$

9.48 (a) $\lambda=-2, u=(1,1,0), v=(1,0,-1)$; $\lambda=4, w=(1,1,2)$,
(b) $\lambda=2, u=(1,1,0)$; $\lambda=-4, v=(0,1,1)$,
(c) $\lambda=3, u=(1,1,0), v=(1,0,1)$; $\lambda=1, w=(2,-1,1)$. Somente A e C podem ser diagonalizadas; usamos $P=[u,v,w]$.

9.49 (a) $\lambda=2, u=(3,-1)$; $\lambda=6, v=(1,1)$, (b) Não tem autovalores reais.

9.50 Precisamos $[-\mathrm{tr}(A)]^2-4[\det(A)]\geq 0$ ou $(a-d)^2+4bc\geq 0$.

9.51 $A=[1,1;\quad 0,1]$

9.52 (a) $P=[2,-1;\quad 1,2]/\sqrt{5}$, $D=[7,0;\quad 0,3]$,
(b) $P=[1,1;\quad 1,-1]/\sqrt{2}$, $D=[3,0;\quad 0,5]$,
(c) $P=[3,-1;\quad 1,3]/\sqrt{10}$, $D=[8,0;\quad 0,2]$

9.57 (a) $\lambda=-1$, $u=(1,-1,0)$, $v=(1,1,-2)$; $\lambda=2$, $w=(1,1,1)$,
(b) $\lambda=1$, $u=(2,1,-1)$, $v=(2,-3,1)$; $\lambda=22$, $w=(1,2,4)$;
Normalizamos u, v, w, obtendo $\hat{u}, \hat{v}, \hat{w}$ e tomamos $P=[\hat{u},\hat{v},\hat{w}]$. (*Observação*: u e v não são únicos.)

9.58 (a) $x=(4s+t)/\sqrt{17}$, $y=(-s+4t)/\sqrt{17}$, $q(s,t)=5s^2-12t^2$,
(b) $x=(3s-t)/\sqrt{10}$, $y=(s+3t)/\sqrt{10}$, $q(s,t)=s^2+11t^2$

9.59 (a) $x=(3s+2t)/\sqrt{13}$, $y=r$, $z=(2s-3t)/\sqrt{13}$, $q(r,s,t)=3r^2+9s^2-4t^2$,
(b) $x=5Ks+Lt$, $y=Jr+2Ks-2Lt$, $z=2Jr-Ks-Lt$, onde $J=1/\sqrt{5}$, $K=1/\sqrt{30}$,
$L=1/\sqrt{6}$; $q(r,s,t)=2r^2+2s^2+8t^2$

9.60 (a) $A=\frac{1}{2}[5,-3;\ -3,5]$, $B=\frac{1}{2}[3,-1;\ -1,3]$,
(b) $A=\frac{1}{5}[14,-2;\ -2,11]$, $B=\frac{1}{5}[\sqrt{2}+4\sqrt{3}, 2\sqrt{2}-2\sqrt{3};\ 2\sqrt{2}-2\sqrt{3}, 4\sqrt{2}+\sqrt{3}]$

9.61 (a) $\Delta(t)=m(t)=(t-2)^2(t-6)$, (b) $\Delta(t)=(t-2)^2(t-6)$, $m(t)=(t-2)(t-6)$

9.62 (a) $\Delta(t)=(t-2)^3(t-7)^2$, $m(t)=(t-2)^2(t-7)$,
(b) $\Delta(t)=(t-3)^5$, $m(t)=(t-3)^3$,
(c) $\Delta(t)=(t-2)^2(t-4)^2(t-5)$, $m(t)=(t-2)(t-4)(t-5)$

9.68 Tomamos a matriz companheira A [Exemplo 9.12(b)] com última coluna: (a) $[-8,-6,5]^T$, (b) $[-4,-7,2,5]^T$

9.69 *Sugestão*: A é uma raiz de $h(t)=f(t)-g(t)$ e $h(t)\equiv 0$ ou o grau de $h(t)$ é menor do que o grau de $f(t)$.

Capítulo 10

Formas Canônicas

10.1 INTRODUÇÃO

Seja T um operador linear de um espaço vetorial de dimensão finita. Vimos no Capítulo 6 que T pode não ter uma representação matricial diagonal. Mesmo assim, ainda é possível "simplificar" a representação matricial de T de várias maneiras. É esse o principal assunto deste capítulo. Em particular, obteremos o teorema da decomposição primária e as formas canônicas triangulares, de Jordan e racional.

Observamos que as formas canônicas triangulares e de Jordan existem para um operador T se, e só se, o polinômio característico $\Delta(t)$ de T possuir todas as suas raízes no corpo base K. Isso sempre ocorre se K for o corpo dos complexos $\mathbf{C}$, mas pode não ser verdade se K for o corpo dos reais $\mathbf{R}$.

Neste capítulo também introduzimos a ideia de *espaço quociente*, uma ferramenta muito poderosa que será utilizada na demonstração da existência das formas canônicas triangular e racional.

10.2 FORMA TRIANGULAR

Seja T um operador linear de um espaço vetorial V de dimensão n. Suponha que T possa ser representado pela matriz triangular

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \dots & \dots \\ & & & a_{nn} \end{bmatrix}$$

Então o polinômio característico $\Delta(t)$ de T é um produto de fatores lineares, a saber,

$$\Delta(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

A recíproca também é verdadeira e constitui um teorema importante (demonstrado no Problema 10.28).

Teorema 10.1 Seja $T: V \to V$ um operador linear cujo polinômio característico fatora em polinômios lineares. Então existe uma base de V na qual T é representado por uma matriz triangular.

Teorema 10.1 (Forma Alternativa) Seja A uma matriz quadrada cujo polinômio característico fatora em polinômios lineares. Então A é semelhante a uma matriz triangular, ou seja, existe uma matriz invertível P tal que $P^{-1}AP$ é triangular.

Dizemos que um operador linear T pode ser trazido à forma triangular se T puder ser representado por alguma matriz triangular. Observe que, nesse caso, os autovalores de T são exatamente os elementos que aparecem na diagonal principal. Vejamos uma aplicação dessa observação.

Exemplo 10.1 Seja A uma matriz quadrada sobre o corpo dos complexos $\mathbf{C}$. Suponha que λ seja um autovalor de A^2. Mostre que $\sqrt{\lambda}$ ou $-\sqrt{\lambda}$ é um autovalor de A.

Pelo Teorema 10.1, A e A^2 são semelhantes, respectivamente, a matrizes triangulares no formato

$$B = \begin{bmatrix} \mu_1 & * & \dots & * \\ & \mu_2 & \dots & * \\ & & \dots & \dots \\ & & & \mu_n \end{bmatrix} \qquad \text{e} \qquad B^2 = \begin{bmatrix} \mu_1^2 & * & \dots & * \\ & \mu_2^2 & \dots & * \\ & & \dots & \dots \\ & & & \mu_n^2 \end{bmatrix}$$

Como matrizes semelhantes possuem os mesmos autovalores, $\lambda = \mu_i^2$ para algum i. Assim, $\mu_i = \sqrt{\lambda}$ ou $\mu_i = -\sqrt{\lambda}$ é um autovalor de A.

10.3 INVARIÂNCIA

Seja $T:V \to V$ um operador linear. Dizemos que um subespaço W de V é *invariante por T* ou, então, *T-invariante* se T levar W em si mesmo, ou seja, se $v \in W$ implicar $T(v) \in W$. Nesse caso, a restrição de T a W define um operador linear de W, de maneira que T induz um operador linear $\hat{T}:W \to W$ definido por $\hat{T}(w) = T(w)$ para cada $w \in W$.

Exemplo 10.2

(a) Seja $T: \mathbf{R}^3 \to \mathbf{R}^3$ o operador linear que gira cada vetor v horizontalmente em torno do eixo z por um ângulo θ (ver Figura 10-1), definido por

$$T(x, y, z) = (x \cos \theta - y \operatorname{sen} \theta, \quad x \operatorname{sen} \theta + y \cos \theta, \quad z)$$

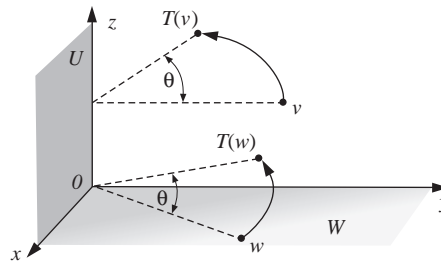



Figura 10-1

Observe que cada vetor $w = (a, b, 0)$ do plano xy, denotado por W, permanece em W aplicando T. Logo, W é T-invariante. Observe também que o eixo z, denotado por U, é invariante por T. Além disso, a restrição de T a W é uma rotação em torno da origem O, ao passo que a restrição de T a U é a transformação identidade de U.

(b) Os autovetores (não nulos) de um operador linear $T:V \to V$ podem ser caracterizados como geradores de subespaços unidimensionais T-invariantes. Suponha que $T(v) = \lambda v$, com $v \neq 0$. Então $W = \{kv: k \in K\}$, o subespaço unidimensional gerado por v, é invariante por T, pois

$$T(kv) = kT(v) = k(\lambda v) = k\lambda v \in W$$

Reciprocamente, suponha que $\dim U = 1$, que $u \neq 0$ gere U e que U seja invariante por T. Então $T(u) \in U$ e, portanto, $T(u)$ é um múltiplo de u, isto é, $T(u) = \mu u$. Logo, u é um autovetor de T.

O teorema seguinte (demonstrado no Problema 10.3) fornece uma classe importante de subespaços invariantes.

Teorema 10.2 Sejam $T:V \to V$ um operador linear e $f(t)$ um polinômio. Então o núcleo de $f(T)$ é invariante por T.

A relação da noção de invariância com as representações matriciais é dada no teorema seguinte (demonstrado no Problema 10.5).

Teorema 10.3 Seja W um subespaço invariante por $T:V \to V$. Então existe uma representação matricial de T com uma matriz em blocos $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, sendo A a representação matricial da restrição $\hat{T}$ de T a W.

10.4 DECOMPOSIÇÃO EM SOMAS DIRETAS INVARIANTES

Dizemos que um espaço vetorial V é a *soma direta* dos subespaços $W_1, W_2, ..., W_r$, denotada por

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_r$$

se cada vetor $v \in V$ puder ser escrito, de maneira única, na forma

$$v = w_1 + w_2 + \ldots + w_r, \text{ com } w_i \in W_i$$

Vale o teorema seguinte (demonstrado no Problema 10.7).

Teorema 10.4 Sejam $W_1, W_2, ..., W_r$ subespaços de V e suponha que

$$B_1 = \{w_{11}, w_{12}, \ldots, w_{1n_1}\}, \qquad \ldots, \qquad B_r = \{w_{r1}, w_{r2}, \ldots, w_{rn_r}\}$$

sejam bases de $W_1, W_2, ..., W_r$, respectivamente. Então V é a soma direta dos W_i se, e só se, a união $B = B_1 \cup \ldots \cup B_r$ é uma base de V.

Suponha, agora, que $T:V \to V$ seja um operador linear e que V seja a soma direta de subespaços (não nulos) T-invariantes $W_1, W_2, ..., W_r$, ou seja,

$$V = W_1 \oplus \ldots \oplus W_r \quad \text{e} \quad T(W_i) \subseteq W_i, \quad i = 1, \ldots, r$$

Denotando a restrição de T a W_i por T_i, dizemos que T é *decomponível* nos operadores T_i ou que T é a *soma direta* dos T_i, e escrevemos $T = T_1 \oplus \ldots \oplus T_r$. Também dizemos que os subespaços $W_1, W_2, ..., W_r$ *reduzem T* ou que formam uma *decomposição em soma direta T-invariante de V*.

Consideremos o caso especial em que dois subespaços U e W reduzem um operador linear $T:V \to V$; digamos, com dim $U = 2$ e dim $W = 3$, sendo $\{u_1, u_2\}$ e $\{w_1, w_2, w_3\}$ bases de U e W, respectivamente. Denotando as restrições de T a U e W por T_1 e T_2, respectivamente, temos

$$\begin{array}{ll} T_1(u_1) = a_{11}u_1 + a_{12}u_2 & T_2(w_1) = b_{11}w_1 + b_{12}w_2 + b_{13}w_3 \\ T_1(u_2) = a_{21}u_1 + a_{22}u_2 & T_2(w_2) = b_{21}w_1 + b_{22}w_2 + b_{23}w_3 \\ & T_2(w_3) = b_{31}w_1 + b_{32}w_2 + b_{33}w_3 \end{array}$$

Por isso, as representações matriciais de T_1, T_2, T são, respectivamente, as matrizes A, B, M dadas por

$$A = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix}, \qquad M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

De fato, a matriz diagonal em blocos M é a representação de T na base $\{u_1, u_2, w_1, w_2, w_3\}$ de V (Teorema 10.4), já que $T(u_i) = T_1(u_i)$ e $T(w_j) = T_2(w_j)$

Uma generalização desse argumento fornece o teorema a seguir.

Teorema 10.5 Seja $T:V \to V$ um operador linear e suponha que V seja a soma direta de subespaços T-invariantes, digamos, $W_1, W_2, ..., W_r$. Se A_i é a representação matricial da restrição de T a W_i então T pode ser representado pela matriz diagonal em blocos

$$M = \text{diag}(A_1, A_2, ..., A_r)$$

10.5 DECOMPOSIÇÃO PRIMÁRIA

O teorema seguinte mostra que qualquer operador linear $T:V \to V$ é decomponível em operadores cujos polinômios mínimos são potências de polinômios irredutíveis. Esse é o primeiro passo na direção de obter uma forma canônica para T.

Teorema 10.6 (Teorema da Decomposição Primária) Seja $T:V \to V$ um operador linear com polinômio mínimo

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_r(t)^{n_r}$$

em que os $f_i(t)$ são polinômios mônicos irredutíveis distintos. Então V é a soma direta dos subespaços T-invariantes $W_1, W_2, ..., W_r$, sendo W_i o núcleo de $f_i(T)^{n_i}$. Além disso, $f_i(t)^{n_i}$ é o polinômio mínimo da restrição de T a W_i.

Os polinômios $f_i(t)^{n_i}$ do teorema são primos entre si. Portanto, esse teorema fundamental é uma consequência direta (Problema 10.11) dos dois teoremas seguintes (demonstrados nos Problemas 10.9 e 10.10, respectivamente).

Teorema 10.7 Sejam $T:V \to V$ um operador linear e $f(t) = g(t)h(t)$ um polinômio tal que $f(T) = 0$, com $g(t)$ e $h(t)$ polinômios primos entre si. Então V é a soma direta dos subespaços T-invariantes $U = \text{Nuc}\, g(T)$ e $W = \text{Nuc}\, h(T)$.

Teorema 10.8 No Teorema 10.7, se $f(t)$ é o polinômio mínimo de T [e $g(t)$ e $h(t)$ são mônicos] então $g(t)$ e $h(t)$ são os polinômios mínimos das restrições de T a U e W, respectivamente.

O teorema da decomposição primária também é utilizado para apresentar uma caracterização bastante útil dos operadores diagonalizáveis (demonstrada no Problema 10.12).

Teorema 10.9 Um operador linear $T:V \to V$ é diagonalizável se, e só se, o polinômio mínimo $m(t)$ de T é um produto de polinômios lineares distintos.

Teorema 10.9 (Forma Alternativa) Uma matriz A é semelhante a uma matriz diagonal se, e só se, o polinômio mínimo de A é um produto de polinômios lineares distintos.

Exemplo 10.3 Seja $A \neq I$ uma matriz quadrada tal que $A^3 = I$. Decida se A é ou não semelhante a uma matriz diagonal sobre (i) o corpo real $\mathbf{R}$, (ii) o corpo complexo $\mathbf{C}$.

Como $A^3 = I$, A é uma raiz do polinômio $f(t) = t^3 - 1 = (t-1)(t^2+t+1)$. O polinômio mínimo $m(t)$ de A não pode ser $t - 1$, porque $A \neq I$. Logo,

$$m(t) = t^2 + t + 1 \qquad \text{ou} \qquad m(t) = t^3 - 1$$

Como nenhum desses polinômios é um produto de fatores lineares sobre $\mathbf{R}$, A não é diagonalizável sobre $\mathbf{R}$. Por outro lado, cada um desses polinômios é um produto de polinômios lineares distintos sobre $\mathbf{C}$. Logo, A é diagonalizável sobre $\mathbf{C}$.

10.6 OPERADORES NILPOTENTES

Dizemos que um operador linear $T:V \to V$ é *nilpotente* se $T^n = 0$, para algum inteiro positivo n. Dizemos que o *índice de nilpotência* de T é k se $T^k = 0$, mas $T^{k-1} \neq 0$. Analogamente, dizemos que uma matriz quadrada A é *nilpotente* se $A^n = 0$ para algum inteiro positivo n, e de índice k se $A^k = 0$, mas $A^{k-1} \neq 0$. Claramente, o polinômio mínimo de um operador (matriz) nilpotente de índice k é $m(t) = t^k$ e, por isso, 0 é seu único autovalor.

Exemplo 10.4 As duas matrizes quadradas de ordem r dadas a seguir são utilizadas em todo este capítulo.

$$N = N(r) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dots\dots\dots\dots\dots\dots} \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad J(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dots\dots\dots\dots\dots\dots} \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}$$

A primeira matriz, denominada *bloco de Jordan nilpotente*, consiste em entradas iguais a 1 na linha acima da diagonal, que denominamos *superdiagonal*, e 0 em todas as demais entradas. $N = N(r)$ é uma matriz nilpotente de ordem r. (A matriz $N(1)$ de ordem 1 é, simplesmente, a matriz nula [0] de ordem 1.)

A segunda matriz, denominada *bloco de Jordan* associado ao autovalor λ, consiste em entradas iguais a λ na diagonal, 1 na superdiagonal e 0 em todas as demais entradas. Observe que

$$J(\lambda) = \lambda I + N$$

De fato, mostraremos que qualquer operador linear T pode ser decomposto em operadores, cada um dos quais é uma soma de um operador escalar com um operador nilpotente.

Um resultado fundamental sobre operadores nilpotentes (demonstrado no Problema 10.16) é o teorema seguinte.

Teorema 10.10 Seja $T: V \to V$ um operador nilpotente de índice k. Então T possui uma representação matricial diagonal em blocos em que cada bloco diagonal é um bloco de Jordan nilpotente N. Há, pelo menos, um bloco N de ordem k e todos os outros blocos N são de ordens $\leq k$. O número de blocos N de cada ordem possível é determinado, de modo único, pelo operador T. O número total de blocos N de todas as ordens é igual à nulidade de T.

A demonstração do Teorema 10.10 esclarece que o número de blocos N de ordem i é igual a $2m_i - m_{i+1} - m_{i-1}$, onde m_i é a nulidade de T^i.

10.7 FORMA CANÔNICA DE JORDAN

Um operador linear T pode ser colocado em forma canônica de Jordan se seus polinômios mínimo e característico podem ser fatorados em polinômios lineares. Isso sempre vale se K for o corpo dos complexos $\mathbf{C}$. Em geral, sempre podemos estender o corpo base K a um corpo no qual os polinômios mínimo e característico podem ser fatorados em polinômios lineares; assim, nesse sentido amplo, qualquer operador tem uma forma canônica de Jordan. Analogamente, toda matriz é semelhante a uma matriz em forma canônica de Jordan.

O teorema seguinte (demonstrado no Problema 10.18) descreve a *forma canônica de Jordan J* de um operador linear T.

Teorema 10.11 Seja $T: V \to V$ um operador linear cujos polinômios mínimo e característico são, respectivamente,

$$\Delta(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \qquad \text{e} \qquad m(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}$$

e em que os λ_i são escalares distintos. Então T tem uma representação matricial diagonal em blocos, em que cada bloco diagonal é um bloco de Jordan $J_{ij} = J(\lambda_i)$. Para cada λ_{ij}, o bloco J_{ij} correspondente tem as propriedades listadas a seguir.

(i) Existe pelo menos um J_{ij} de ordem m_i e todos os outros J_{ij} são de ordem $\leq m_i$.
(ii) A soma das ordens dos J_{ij} é n_i.
(iii) O número de J_{ij} é a multiplicidade geométrica de λ_i.
(iv) O número de J_{ij} de cada ordem possível é determinado, de modo único, pelo operador T.

Exemplo 10.5 Suponha que os polinômios mínimo e característico de um operador linear T sejam, respectivamente,

$$\Delta(t) = (t-2)^4(t-5)^3 \qquad \text{e} \qquad m(t) = (t-2)^2(t-5)^3$$

Então a forma canônica de Jordan de T é uma das matrizes diagonais em blocos a seguir.

$$\operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}\right) \quad \text{ou} \quad \operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2], [2], \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}\right)$$

Se T tiver dois autovetores independentes associados ao autovalor 2, então a forma canônica de Jordan é a primeira matriz; a segunda ocorre se T tiver três autovetores independentes associados a 2.

10.8 SUBESPAÇOS CÍCLICOS

Seja T um operador linear de um espaço vetorial V de dimensão finita sobre K. Suponha que $v \in V$, com $v \neq 0$. O conjunto de todos os vetores da forma $f(T)v$, em que $f(t)$ percorre todos os polinômios sobre K, é um subespaço T-invariante de V denominado *subespaço T-cíclico de V gerado por* v e denotado por $Z(v, T)$. Pelo Problema 10.56, poderíamos ter definido, equivalentemente, $Z(v, T)$ como a interseção de todos os subespaços T-invariantes de V que contêm v. A restrição de T a $Z(v, T)$ é denotada por T_v.

Consideremos, agora, a sequência

$$v, \quad T(v), \quad T^2(v), \quad T^3(v), \quad \ldots$$

das potências de T agindo sobre v. Seja k o menor inteiro tal que $T^k(v)$ é uma combinação linear daqueles vetores que o precedem nessa sequência,digamos,

$$T^k(v) = -a_{k-1}T^{k-1}(v) - \cdots - a_1T(v) - a_0v$$

Então

$$m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1t + a_0$$

é o único polinômio mônico de menor grau tal que $m_v(T)(v) = 0$. Dizemos que $m_v(t)$ é o *T-anulador de* v *e de* $Z(v, T)$.

Vale o teorema seguinte (demonstrado no Problema 10.29), em que usamos a notação que acabamos de introduzir.

Teorema 10.12 Valem as afirmações seguintes.

(i) O conjunto set $\{v, T(v), \ldots, T^{k-1}(v)\}$ é uma base de $Z(v, T)$, portanto, dim $Z(v, T) = k$.

(ii) O polinômio mínimo de T_v é $m_v(t)$.

(iii) A representação matricial de T_v na base de (i) é a *matriz companheira* $C(m_v)$ de $m_v(t)$, ou seja,

$$C(m_v) = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \ldots & 1 & -a_{k-1} \end{bmatrix}$$

10.9 FORMA CANÔNICA RACIONAL

Nesta seção apresentamos a forma canônica racional de um operador linear $T{:}V \to V$. Enfatizamos que essa forma existe sempre, mesmo que o polinômio mínimo não possa ser fatorado em polinômios lineares. (Isso não ocorre com a forma canônica de Jordan.)

Lema 10.13 Seja $T:V \to V$ um operador linear cujo polinômio mínimo é $f(t)^n$, em que $f(t)$ é um polinômio mônico irredutível. Então V é a soma direta

$$V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$$

dos subespaços T-cíclicos $Z(v_i, T)$ com T-anuladores correspondentes

$$f(t)^{n_1}, \quad f(t)^{n_2}, \quad \ldots, \quad f(t)^{n_r}, \qquad n = n_1 \geq n_2 \geq \ldots \geq n_r$$

Qualquer outra decomposição de V em subespaços T-cíclicos tem o mesmo número de componentes e o mesmo conjunto de T-anuladores.

Enfatizamos que esse lema (demonstrado no Problema 10.31) não diz que os vetores v_i ou outros subespaços T-cíclicos $Z(v_i, T)$ sejam determinados de modo único por T, mas o lema nos diz que o conjunto dos T-anuladores é determinado, de modo único, por T. Assim, T tem uma única representação matricial diagonal em blocos

$$M = \operatorname{diag}(C_1, C_2, \ldots, C_r)$$

em que as C_i são matrizes companheiras. De fato, as C_i são as matrizes companheiras dos polinômios $f(t)^{n_i}$.

Usando o teorema da decomposição primária e o Lema 10.13, obtemos o resultado a seguir.

Teorema 10.14 Seja $T:V \to V$ um operador linear com polinômio mínimo

$$m(t) = f_1(t)^{m_1} f_2(t)^{m_2} \cdots f_s(t)^{m_s}$$

em que os $f_i(t)$ são polinômios mônicos irredutíveis distintos. Então T tem uma única representação matricial diagonal em blocos

$$M = \operatorname{diag}(C_{11}, C_{12}, \ldots, C_{1r_1}, \ldots, C_{s1}, C_{s2}, \ldots, C_{sr_s})$$

em que as C_{ij} são matrizes companheiras. Em particular, Cij é a matriz companheira do polinômio $f_i(t)^{n_{ij}}$, com

$$m_1 = n_{11} \geq n_{12} \geq \cdots \geq n_{1r_1}, \qquad \ldots, \qquad m_s = n_{s1} \geq n_{s2} \geq \cdots \geq n_{sr_s}$$

A representação matricial de T dada nesse teorema é denominada *forma canônica racional*. Os polinômios $f_i(t)^{n_{ij}}$ são denominados *divisores elementares* de T.

Exemplo 10.6 Seja V um espaço vetorial de dimensão 8 sobre o corpo dos racionais $\mathbf{Q}$, e seja T um operador linear de V cujo polinômio mínimo é

$$m(t) = f_1(t) f_2(t)^2 = (t^4 - 4t^3 + 6t^2 - 4t - 7)(t - 3)^2$$

Assim, como $\dim V = 8$, o polinômio característico de T é $\Delta(t) = f_1(t) f_2(t)^4$. Também vemos que a forma canônica racional M de T precisa ter um bloco dado pela matriz companheira de $f_1(t)$ e um bloco dado pela matriz companheira de $f_2(t)^2$. Temos duas possibilidades.

(a) $\operatorname{diag}[C(t^4 - 4t^3 + 6t^2 - 4t - 7), \quad C((t-3)^2), \quad C((t-3)^2)]$

(b) $\operatorname{diag}[C(t^4 - 4t^3 + 6t^2 - 4t - 7), \quad C((t-3)^2), \quad C(t-3), C(t-3)]$

Ou seja,

$$\text{(a)} \quad \operatorname{diag}\left(\begin{bmatrix} 0 & 0 & 0 & 7 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & -9 \\ 1 & 6 \end{bmatrix}, \begin{bmatrix} 0 & -9 \\ 1 & 6 \end{bmatrix} \right), \text{(b)} \quad \operatorname{diag}\left(\begin{bmatrix} 0 & 0 & 0 & 7 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & -9 \\ 1 & 6 \end{bmatrix}, [3], [3] \right)$$

10.10 ESPAÇO QUOCIENTE

Sejam V um espaço vetorial sobre um corpo K e W um subespaço de V. Se v é um vetor qualquer de V, escrevemos $v + W$ para o conjunto das somas $v + w$, com $w \in W$; ou seja,

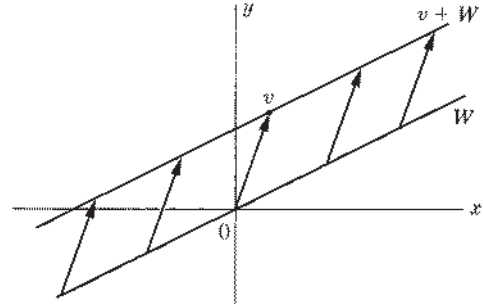
$$v + W = \{v + w : w \in W\}$$

Esses conjuntos são denominados *classes laterais* de W em V. Mostramos (Problema 10.22) que essas classes laterais definem uma partição de V em subconjuntos dois a dois disjuntos.

Figura 10-2

Exemplo 10.7 Seja W o subespaço de $\mathbf{R}^2$ definido por

$$W = \{(a, b) : a = b\},$$

ou seja, W é a reta dada pela equação $x - y = 0$. Podemos ver $v + W$ como uma translação dessa reta obtida com a soma do vetor v a cada ponto de W. Conforme indicado na Figura 10-2, a classe lateral $v + W$ também é uma reta, e é paralela a W. Assim, as classes laterais de W em $\mathbf{R}^2$ são precisamente todas as retas paralelas a W.

No teorema seguinte usamos as classes laterais de um subespaço W de um espaço vetorial V para definir um novo espaço vetorial, denominado *espaço quociente* de V por W e denotado por V/W.

Teorema 10.15 Seja W um subespaço de um espaço vetorial V sobre um corpo K. Então as classes laterais de W em V formam um espaço vetorial sobre K com as operações de adição e multiplicação por escalar dadas a seguir.

$$\text{(i)}\ (u + w) + (v + W) = (u + v) + W, \text{(ii)}\ r(u + W) = ru + W, \text{com } r \in K$$

Observe que, na demonstração do Teorema 10.15 (Problema 10.24), é necessário mostrar, antes de mais nada, que as operações de V/W estão bem definidas, ou seja, sempre que $u + W = u' + W$ e $v + W = v' + W$, temos

$$\text{(i)}\ (u + v) + W = (u' + v') + W \quad \text{e} \quad \text{(ii)}\ ru + W = ru' + W, \text{com } r \in K.$$

No caso de um subespaço invariante, temos o seguinte resultado útil (demonstrado no Problema 10.27).

Teorema 10.16 Seja W um subespaço invariante por um operador linear $T: V \to V$. Então T induz um operador linear $\bar{T}$ de V/W definido por $\bar{T}(v + W) = T(v) + W$. Além disso, se T é uma raiz de algum polinômio, então $\bar{T}$ também é. Assim, o polinômio mínimo de $\bar{T}$ divide o polinômio mínimo de T.

Problemas Resolvidos

Subespaços invariantes

10.1 Seja $T: V \to V$ um operador linear. Mostre que cada um dos subespaços dados é invariante por T.

(a) $\{0\}$, (b) V, (c) o núcleo de T, (d) a imagem de T.

(a) Temos $T(0) = 0 \in \{0\}$; logo, $\{0\}$ é invariante por T.

(b) Para cada $v \in V$, $T(v) \in V$; logo, V é invariante por T.

(c) Seja $u \in \text{Nuc}\, T$. Então $T(u) = 0 \in \text{Nuc}\, T$, porque o núcleo de T é um subespaço de V. Assim, Nuc T é invariante por T.

(d) Como $T(v) \in \text{Im}\, T$ para cada $v \in V$, isso certamente vale quando $v \in \text{Im}\, T$. Logo, a imagem de T é invariante por T.

10.2 Seja $\{W_i\}$ uma coleção de subespaços T-invariantes de um espaço vetorial V. Mostre que a interseção $W = \bigcap_i W_i$ também é T-invariante.

Se $v \in W$, então $v \in W_i$ para cada i. Como W_i é T-invariante, temos $T(v) \in W_i$, para cada i. Assim, $T(v) \in W$ e, portanto, W é T-invariante.

10.3 Demonstre o Teorema 10.2. Sejam $T:V \to V$ um operador linear e $f(t)$ um polinômio. Então o núcleo de $f(T)$ é invariante por T.

Suponha que $v \in \text{Nuc}\, f(T)$, ou seja, que $f(T)(v) = 0$. Devemos mostrar que $T(v)$ também pertence ao núcleo de $f(T)$, ou seja, que $f(T)(T(v)) = (f(T) \circ T)(v) = 0$. Como $f(t)t = tf(t)$, temos $f(T) \circ T = T \circ f(T)$. Assim, como queríamos, resulta

$$(f(T) \circ T)(v) = (T \circ f(T))(v) = T(f(T)(v)) = T(0) = 0$$

10.4 Encontre todos subespaços invariantes por $A = \begin{bmatrix} 2 & -5 \\ 1 & -2 \end{bmatrix}$ interpretada como um operador de $\mathbf{R}^2$.

Pelo Problema 10.1, $\mathbf{R}^2$ e $\{0\}$ são invariantes por A. Agora, se A tiver algum outro subespaço invariante, deve ser unidimensional. Contudo, o polinômio característico de A é

$$\Delta(t) = t^2 - \text{tr}(A)\, t + |A| = t^2 + 1$$

Logo, A não tem autovalores (em $\mathbf{R}$) e, portanto, A não tem autovetores. Como os subespaços unidimensionais correspondem a autovetores [Exemplo 10.2(b)], segue que $\mathbf{R}^2$ e $\{0\}$ são os únicos subespaços invariantes por A.

10.5 Demonstre o Teorema 10.3. Seja W um subespaço T-invariante. Então existe uma representação matricial de T com uma matriz em blocos $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, sendo A a representação matricial da restrição $\hat{T}$ de T a W.

Escolhemos uma base $\{w_1, ..., w_r\}$ de W e a estendemos a uma base $\{w_1, \ldots, w_r, v_1, \ldots, v_s\}$ de V. Temos

$$\begin{aligned}
\hat{T}(w_1) &= T(w_1) = a_{11}w_1 + \cdots + a_{1r}w_r \\
\hat{T}(w_2) &= T(w_2) = a_{21}w_1 + \cdots + a_{2r}w_r \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\hat{T}(w_r) &= T(w_r) = a_{r1}w_1 + \cdots + a_{rr}w_r \\
T(v_1) &= b_{11}w_1 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s \\
T(v_2) &= b_{21}w_1 + \cdots + b_{2r}w_r + c_{21}v_1 + \cdots + c_{2s}v_s \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(v_s) &= b_{s1}w_1 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}v_s
\end{aligned}$$

Como a matriz de T nessa base é a transposta da matriz de coeficientes desse sistema de equações (Seção 6.2), decorre que essa matriz é da forma $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, sendo A a transposta da matriz de coeficientes do subsistema óbvio. Pelo mesmo argumento, A é a matriz de $\hat{T}$ em relação à base $\{w_i\}$ de W.

10.6 Denote por $\hat{T}$ a restrição de um operador T a um subespaço T-invariante W. Prove as afirmações dadas.

(a) Para cada polinômio $f(t)$, vale $f(\hat{T})(w) = f(T)(w)$.

(b) O polinômio mínimo de $\hat{T}$ divide o polinômio mínimo de T.

(a) Se $f(t) = 0$ ou se $f(t)$ é uma constante (isto é, de grau 1), o resultado é imediato.

Suponha que o grau de $f(t)$ seja $n > 1$ e que o resultado valha para polinômios de grau menor do que n. Seja

$$f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Então

$$\begin{aligned}
f(\hat{T})(w) &= (a_n\hat{T}^n + a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) \\
&= (a_n\hat{T}^{n-1})(\hat{T}(w)) + (a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) \\
&= (a_n T^{n-1})(T(w)) + (a_{n-1}T^{n-1} + \cdots + a_0 I)(w) = f(T)(w)
\end{aligned}$$

(b) Seja $m(t)$ o polinômio mínimo de T. Então, por (a), $m(\hat{T})(w) = m(T)(w) = 0(w) = 0$ para cada $w \in W$; ou seja, $\hat{T}$ é um zero do polinômio $m(t)$. Assim, concluímos que o polinômio mínimo de $\hat{T}$ divide $m(t)$.

Decomposições em somas diretas invariantes

10.7 Demonstre o Teorema 10.4. Sejam W_1, W_2, ..., W_r subespaços de V com respectivas bases

$$B_1 = \{w_{11}, w_{12}, \ldots, w_{1n_1}\}, \qquad \ldots, \qquad B_r = \{w_{r1}, w_{r2}, \ldots, w_{rn_r}\}$$

Então V é a soma direta dos W_i se, e só se, a união $B = \bigcup_i B_i$ é uma base de V.

Seja B uma base de V. Então, para cada $v \in V$,

$$v = a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} \;+\; \cdots \;+\; a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = w_1 + w_2 + \cdots + w_r$$

com $w_i = a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} \in W_i$. Agora mostramos que somas como essa são únicas. Suponha que

$$v = w'_1 + w'_2 + \cdots + w'_r, \qquad \text{com} \qquad w'_i \in W_i$$

Como $\{w_{i1}, \ldots, w_{in_i}\}$ é uma base de W_i, temos W_i, $w'_i = b_{i1}w_{i1} + \cdots + b_{in_i}w_{in_i}$, e, portanto,

$$v = b_{11}w_{11} + \cdots + b_{1n_1}w_{1n_1} \;+\; \cdots \;+\; b_{r1}w_{r1} + \cdots + b_{rn_r}w_{rn_r}$$

Como B é uma base de V, temos $a_{ij} = b_{ij}$, para cada i e cada j. Logo, $w_i = w'_i$, e, portanto, a soma de v é única. Em vista disso, V é a soma direta dos W_i.

Reciprocamente, suponha que V seja a soma direta dos W_i. Então, para cada $v \in V$, $v = w_1 + \cdots + w_r$, onde $w_i \in W_i$. Como $\{w_{ij_i}\}$ é uma base de W_i, cada w_i é uma combinação linear dos w_{ij_i}, e, portanto, v é uma combinação linear dos elementos de B. Assim, B gera V. Mostremos que B é linearmente independente. Suponha que

$$a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} \;+\; \cdots \;+\; a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = 0$$

Observe que $a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} \in W_i$. Também temos $0 = 0 + 0 \cdots 0 \in W_i$. Como uma soma dessas para 0 é única,

$$a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} = 0 \qquad \text{para} \qquad i = 1, \ldots, r$$

A independência das bases $\{w_{ij_i}\}$ implica que todos os coeficientes a são nulos. Assim, B é linearmente independente e, portanto, uma base de V.

10.8 Seja $T{:}V \to V$ um operador linear e suponha que $T = T_1 \oplus T_2$ em relação a uma decomposição em soma direta T-invariante $V = U \oplus W$. Prove as afirmações dadas.

(a) $m(t)$ é o mínimo múltiplo comum de $m_1(t)$ e $m_2(t)$, sendo $m(t)$, $m_1(t)$, $m_2(t)$ os polinômios mínimos de T, T_1, T_2, respectivamente.

(b) $\Delta(t) = \Delta_1(t)\Delta_2(t)$, sendo $\Delta(t), \Delta_1(t), \Delta_2(t)$ os polinômios característicos de T, T_1, T_2, respectivamente.

(a) Pelo Problema 10.6, ambos $m_1(t)$ e $m_2(t)$ dividem $m(t)$. Suponha, agora, que $f(t)$ seja um múltiplo de ambos $m_1(t)$ e $m_2(t)$. Então $f(T_1)(U) = 0$ e $f(T_2)(W) = 0$. Seja $v \in V$, de modo que $v = u + w$ com $u \in U$ e $w \in W$. Agora,

$$f(T)v = f(T)u + f(T)w = f(T_1)u + f(T_2)w = 0 + 0 = 0$$

Ou seja, T é um zero de $f(t)$. Logo, $m(t)$ divide $f(t)$, de modo que $m(t)$ é o mínimo múltiplo comum de $m_1(t)$ e $m_2(t)$.

(b) Pelo Teorema 10.5, T tem uma representação matricial $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, em que A e B são as representações matriciais de T_1 e T_2, respectivamente. Então, como queríamos,

$$\Delta(t) = |tI - M| = \begin{vmatrix} tI - A & 0 \\ 0 & tI - B \end{vmatrix} = |tI - A||tI - B| = \Delta_1(t)\Delta_2(t)$$

10.9 Demonstre o Teorema 10.7. Sejam $T{:}V \to V$ um operador linear e $f(t) = g(t)h(t)$ um polinômio tal que $f(T) = 0$, com $g(t)$ e $h(t)$ polinômios primos entre si. Então V é a soma direta dos subespaços T-invariantes $U = \text{Nuc}\, g(T)$ e $W = \text{Nuc}\, h(T)$.

Inicialmente, observe que U e W são T-invariantes pelo Teorema 10.2. Agora, sendo $g(t)$ e $h(t)$ polinômios primos entre si, existem polinômios $r(t)$ e $s(t)$ tais que

$$r(t)g(t) + s(t)h(t) = 1$$

Logo, para o operador T,
$$r(T)g(T) + s(T)h(T) = I \qquad (*)$$

Seja $v \in V$; então, por (*),
$$v = r(T)g(T)v + s(T)h(T)v$$

Mas a primeira parcela dessa soma pertence a $W = \text{Nuc}\, h(T)$, porque

$$h(T)r(T)g(T)v = r(T)g(T)h(T)v = r(T)f(T)v = r(T)0v = 0$$

Analogamente, a segunda parcela pertence a U. Assim, V é a soma de U e W.
Para provar que $V = U \oplus W$, devemos mostrar que uma soma $v = u + w$, com $u \in U$ e $w \in W$, é determinada de modo único por v. Aplicando o operador $r(T)g(T)$ a $v = u + w$, e usando que $g(T)u = 0$, obtemos

$$r(T)g(T)v = r(T)g(T)u + r(T)g(T)w = r(T)g(T)w$$

Também, aplicando (*) unicamente a w e usando que $h(T)w = 0$, obtemos

$$w = r(T)g(T)w + s(T)h(T)w = r(T)g(T)w$$

Ambas fórmulas juntas nos dão $w = r(T)g(T)v$ e, portanto, w está determinado de modo único por v. Analogamente, u está determinado de modo único por v. Assim, $V = U \oplus W$, como queríamos mostrar.

10.10 Demonstre o Teorema 10.8. No Teorema 10.7 (Problema 10.9), se $f(t)$ é o polinômio mínimo de T (e $g(t)$ e $h(t)$ são mônicos) então $g(t)$ e $h(t)$ são os polinômios mínimos das restrições T_1 e T_2 de T a U e W, respectivamente.

Sejam $m_1(t)$ e $m_2(t)$ os polinômios mínimos de T_1 e T_2, respectivamente. Observe que $g(T_1) = 0$ e $h(T_2) = 0$, porque $U = \text{Nuc}\, g(T)$ e $W = \text{Nuc}\, h(T)$. Assim,

$$m_1(t) \text{ divide } g(t) \qquad \text{e} \qquad m_2(t) \text{ divide } h(t) \qquad (1)$$

Pelo Problema 10.9, $f(t)$ é o mínimo múltiplo comum de $m_1(t)$ e $m_2(t)$. No entanto, $m_1(t)$ e $m_2(t)$ são primos entre si, porque $g(t)$ e $h(t)$ são primos entre si. Em vista disso, $f(t) = m_1(t)m_2(t)$. Também temos $f(t) = g(t)h(t)$. Essas duas equações, junto com (1), e lembrando que todos polinômios envolvidos são mônicos, implicam que $g(t) = m_1(t)$ e $h(t) = m_2(t)$, como queríamos mostrar.

10.11 Demonstre o Teorema da Decomposição Primária 10.6. Seja $T{:}V \to V$ um operador linear com polinômio mínimo

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \dots f_r(t)^{n_r}$$

em que os $f_i(t)$ são polinômios mônicos irredutíveis distintos. Então V é a soma direta dos subespaços T-invariantes $W_1, ..., W_r$, sendo W_i o núcleo de $f_i(T)^{n_i}$. Além disso, $f_i(t)^{n_i}$ é o polinômio mínimo da restrição de T a W_i.

A prova é por indução em r. O caso $r = 1$ é trivial. Suponha que o teorema tenha sido provado para $r - 1$. Pelo Teorema 10.7, podemos escrever V como a soma direta de subespaços T-invariantes W_1 e V_1, em que W_1 é o núcleo de $f_1(T)^{n_1}$ e V_1 é o núcleo de $f_2(T)^{n_2} \cdots f_r(T)^{n_r}$. Pelo Teorema 10.8, os polinômios mínimos das restrições de T a W_1 e V_1 são $f_1(t)^{n_1}$ e $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$, respectivamente.

Denotemos a restrição de T a V_1 por $\hat{T}_1$. Pela hipótese de indução, V_1 é a soma direta de subespaços $W_2, ..., W_r$ tais que W_i é o núcleo de $f_i(T_1)^{n_i}$ e tais que $f_i(t)^{n_i}$ é o polinômio mínimo da restrição de $\hat{T}_1$ a W_i. Ocorre que o núcleo de $f_i(T)^{n_i}$, com $i = 2, .., r$, necessariamente está contido em V_1, pois $f_i(t)^{n_i}$ divide $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$. Assim, o núcleo de $f_i(T)^{n_i}$ é igual ao núcleo de $f_i(T_1)^{n_i}$, que é W_i. Também, a restrição de T a W_i é igual à restrição de $\hat{T}_1$ a W_i (com $i = 2, ..., r$). Logo, $f_i(t)^{n_i}$ é também o polinômio mínimo da restrição de T a W_i. Assim, $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ é a decomposição procurada de T.

10.12 Demonstre o Teorema 10.9. Um operador linear $T{:}V \to V$ tem uma representação matricial diagonal se, e só se, o polinômio mínimo $m(t)$ de T é um produto de polinômios lineares distintos.

Digamos que $m(t)$ seja um produto de polinômios lineares distintos

$$m(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_r)$$

em que os λ_i são escalares distintos. Pelo teorema da decomposição primária, V é a soma direta de subespaços W_1, ..., W_r, em que $W_i = \text{Nuc}\,(T - \lambda_i I)$. Assim, se $v \in W_i$, então $(T - \lambda_i I)(v) = 0$, ou $T(v) = \lambda_i v$. Em outras palavras, cada vetor de W_i é um autovetor associado ao autovalor λ_i. Pelo Teorema 10.4, a união de bases dos subespaços W_1, ..., W_r é uma base de V. Como essa base consiste em autovetores de T, concluímos que T é diagonalizável.

Reciprocamente, digamos que T seja diagonalizável (isto é, que V tenha uma base consistindo de autovetores de T). Sejam $\lambda_1, \ldots, \lambda_s$ os autovalores distintos de T. Então o operador

$$f(T) = (T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_s I)$$

leva cada vetor da base em 0. Assim, $f(T) = 0$ e, portanto, o polinômio mínimo $m(t)$ de T divide o polinômio

$$f(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_s I)$$

Em vista disso, $m(t)$ é um produto de polinômios lineares distintos.

Operadores nilpotentes, forma canônica de Jordan

10.13 Seja $T: V \to V$ um operador linear. Suponha que, para $v \in V$, tenhamos $T^k(v) = 0$, mas $T^{k-1}(v) \neq 0$. Prove as afirmações dadas.

(a) O conjunto $S = \{v, T(v), \ldots, T^{k-1}(v)\}$ é linearmente independente.

(b) O subespaço W gerado por S é T-invariante.

(c) A restrição $\hat{T}$ de T a W é nilpotente de índice k.

(d) Em relação à base $\{T^{k-1}(v), \ldots, T(v), v\}$ de W, a matriz de T é uma matriz quadrada de ordem k dada pelo bloco de Jordan nilpotente N_k de índice k.

(a) Suponha que

$$av + a_1 T(v) + a_2 T^2(v) + \cdots + a_{k-1} T^{k-1}(v) = 0 \qquad (*)$$

Aplicando T^{k-1} a (*) e usando $T^k(v) = 0$, obtemos $aT^{k-1}(v) = 0$; e, como $T^{k-1}(v) \neq 0$, também $a = 0$. Agora, aplicando T^{k-2} a (*) e usando $T^k(v) = 0$ e $a = 0$, obtemos $a_1 T^{k-1}(v) = 0$; e, portanto, $a_1 = 0$. Em seguida, aplicando T^{k-3} a (*) e usando $T^k(v) = 0$ e $a = a_1 = 0$, obtemos $a_2 T^{k-1}(v) = 0$; e, portanto, $a_2 = 0$. Continuando esse processo, mostramos que todos os a são nulos. Assim, S é independente.

(b) Seja $v \in W$. Então

$$v = bv + b_1 T(v) + b_2 T^2(v) + \cdots + b_{k-1} T^{k-1}(v)$$

Usando $T^k(v) = 0$, temos

$$T(v) = bT(v) + b_1 T^2(v) + \cdots + b_{k-2} T^{k-1}(v) \in W$$

Assim, W é T-invariante.

(c) Por hipótese, $T^k(v) = 0$. Logo, para $i = 0$, ..., $k - 1$,

$$\hat{T}^k(T^i(v)) = T^{k+i}(v) = 0$$

Isso significa que, aplicando $\hat{T}^k$ a cada gerador de W, obtemos 0. Logo, $\hat{T}^k = 0$ e, portanto, $\hat{T}$ é nilpotente de índice k, no máximo. Por outro lado, $\hat{T}^{k-1}(v) = T^{k-1}(v) \neq 0$; ou seja, T é nilpotente de índice exatamente k.

(d) Com a base $\{T^{k-1}(v),\ T^{k-2}(v), \ldots, T(v), v\}$ de W temos

$$\begin{array}{lcl}
\hat{T}(T^{k-1}(v)) & = & T^k(v) = 0 \\
\hat{T}(T^{k-2}(v)) & = & T^{k-1}(v) \\
\hat{T}(T^{k-3}(v)) & = & T^{k-2}(v) \\
\multicolumn{3}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
\hat{T}(T(v)) & = & T^2(v) \\
\hat{T}(v) & = & T(v)
\end{array}$$

Assim, como queríamos mostrar, a matriz de T nessa base é uma matriz quadrada de ordem k dada pelo bloco de Jordan nilpotente N_k.

10.14 Seja $T:V \to V$ um operador linear e denotemos $U = \text{Nuc}\, T^i$ e $W = \text{Nuc}\, T^{i+1}$. Mostre que

(a) $U \subseteq W$, (b) $T(W) \subseteq U$.

(a) Seja $u \in U = \text{Nuc}\, T^i$. Então $T^i(u) = 0$ e, portanto, $T^{i+1}(u) = T(T^i(u)) = T(0) = 0$. Assim, $u \in \text{Nuc}\, T^{i+1} = W$. Como isso vale para cada $u \in U$; resulta $U \subseteq W$.

(b) Analogamente, se $w \in W = \text{Nuc}\, T^{i+1}$, então $T^{i+1}(w) = 0$ e, portanto, $T^{i+1}(w) = T^i(T(w)) = T^i(0) = 0$. Assim, $T(W) \subseteq U$.

10.15 Seja $T:V \to V$ um operador linear e denotemos $X = \text{Nuc}\, T^{i-2}$, $Y = \text{Nuc}\, T^{i-1}$ e $Z = \text{Nuc}\, T^i$. Pelo Problema 10.14, temos $X \subseteq Y \subseteq Z$. Suponha que

$$\{u_1, \ldots, u_r\}, \qquad \{u_1, \ldots, u_r, v_1, \ldots, v_s\}, \qquad \{u_1, \ldots, u_r, v_1, \ldots, v_s, w_1, \ldots, w_t\}$$

sejam bases de X, Y, Z, respectivamente. Mostre que

$$S = \{u_1, \ldots, u_r, T(w_1), \ldots, T(w_t)\}$$

está contido em Y e é linearmente independente.

Pelo Problema 10.14, $T(Z) \subseteq Y$, e, portanto, $S \subseteq Y$. Suponha que S seja linearmente dependente. Então existe uma relação

$$a_1u_1 + \cdots + a_ru_r + b_1T(w_1) + \cdots + b_tT(w_t) = 0$$

em que pelo menos um coeficiente não é nulo. Mais que isso, como $\{u_i\}$ é independente, pelo menos um dos b_k deve ser não nulo. Transpondo, obtemos

$$b_1T(w_1) + \cdots + b_tT(w_t) = -a_1u_1 - \cdots - a_ru_r \in X = \text{Nuc}\, T^{i-2}$$

Logo, $$T^{i-2}(b_1T(w_1) + \cdots + b_tT(w_t)) = 0$$

Assim, $$T^{i-1}(b_1w_1 + \cdots + b_tw_t) = 0, \quad \text{e, portanto,} \quad b_1w_1 + \cdots + b_tw_t \in Y = \text{Nuc}\, T^{i-1}$$

Como $\{u_i, v_j\}$ gera Y, obtemos uma relação entre os u_i, v_j, w_k em que um dos coeficientes (a saber, um dos b_k) não é nulo. Isso contradiz o fato de que $\{u_i, v_j, w_k\}$ é independente. Assim, S também deve ser independente.

10.16 Demonstre o Teorema 10.10. Seja $T:V \to V$ um operador nilpotente de índice k. Então T possui uma representação matricial diagonal em blocos consistindo em blocos de Jordan nilpotente N. Há pelo menos um bloco N de ordem k e todos os outros blocos N são de ordens $\leq k$. O número de blocos N de cada ordem possível é determinado, de modo único, pelo operador T. O número total de blocos N de todas as ordens é igual à nulidade de T.

Suponha que $\dim V = n$ e denotemos $W_1 = \text{Nuc}\, T$, $W_2 = \text{Nuc}\, T^2$, ..., $W_k = \text{Nuc}\, T^k$. Denotemos, também, $m_i = \dim W_i$, para $i = 1, \ldots, k$. Como o índice de T é k, temos $W_k = V$, mas $W_{k-1} \neq V$ e, então, $m_{k-1} < m_k = n$. Pelo Problema 10.14, obtemos

$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_k = V$$

Assim, por indução, podemos escolher uma base $\{u_1, \ldots, u_n\}$ de V tal que $\{u_1, \ldots, u_{m_i}\}$ seja uma base de W_i.

Agora escolhemos uma nova base de V em relação à qual T terá a forma procurada. É conveniente denotar os elementos dessa nova base usando dois índices. Começamos tomando

$$v(1,k) = u_{m_{k-1}+1}, \qquad v(2,k) = u_{m_{k-1}+2}, \qquad \ldots, \qquad v(m_k - m_{k-1}, k) = u_{m_k}$$

bem como

$$v(1,k-1) = Tv(1,k), \qquad v(2,k-1) = Tv(2,k), \qquad \ldots, \qquad v(m_k - m_{k-1}, k-1) = Tv(m_k - m_{k-1}, k)$$

Pelo problema precedente,

$$S_1 = \{u_1 \ldots, u_{m_{k-2}}, v(1,k-1), \ldots, v(m_k - m_{k-1}, k-1)\}$$

é um subconjunto linearmente independente de W_{k-1}. Estendemos S_1 a uma base de W_{k-1} juntando novos elementos (se necessário) que denotamos por

$$v(m_k - m_{k-1} + 1,\ k-1), \qquad v(m_k - m_{k-1} + 2,\ k-1), \qquad \ldots, \qquad v(m_{k-1} - m_{k-2},\ k-1)$$

Em seguida, definimos

$$v(1,k-2) = Tv(1,k-1), \qquad v(2,k-2) = Tv(2,k-1), \qquad \ldots,$$
$$v(m_{k-1} - m_{k-2}, k-2) = Tv(m_{k-1} - m_{k-2}, k-1)$$

Novamente, pelo problema precedente,

$$S_2 = \{u_1, \ldots, u_{m_{k-s}}, v(1, k-2), \ldots, v(m_{k-1} - m_{k-2}, k-2)\}$$

é um subconjunto linearmente independente de W_{k-2}, que pode ser estendido a uma base de W_{k-2} juntando os elementos

$$v(m_{k-1} - m_{k-2} + 1, k-2), \qquad v(m_{k-1} - m_{k-2} + 2, k-2), \qquad \ldots, \qquad v(m_{k-2} - m_{k-3}, k-2)$$

Continuando dessa forma, obtemos uma nova base de V que, para melhor visualização, arranjamos como segue.

$$\begin{array}{lllll}
v(1,k) & \ldots, v(m_k - m_{k-1}, k) & & & \\
v(1,k-1), & \ldots, v(m_k - m_{k-1}, k-1) & \ldots, v(m_{k-1} - m_{k-2}, k-1) & & \\
 & \cdots\cdots\cdots\cdots & & & \\
v(1,2), & \ldots, v(m_k - m_{k-1}, 2), & \ldots, v(m_{k-1} - m_{k-2}, 2), & \ldots, v(m_2 - m_1, 2) & \\
v(1,1), & \ldots, v(m_k - m_{k-1}, 1), & \ldots, v(m_{k-1} - m_{k-2}, 1), & \ldots, v(m_2 - m_1, 1), & \ldots, v(m_1, 1)
\end{array}$$

A linha de baixo forma uma base de W_1, as duas últimas linhas formam uma base de W_2, e assim por diante. No entanto, o que é importante é que T leva cada vetor no vetor que está imediatamente abaixo dele nessa tabela, ou no vetor nulo, se o vetor estiver na linha de baixo da tabela. Assim,

$$T v(i,j) = \begin{cases} v(i, j-1) & \text{se } j > 1 \\ 0 & \text{se } j = 1 \end{cases}$$

Agora fica claro [ver Problema 10.13(d)] que T tem a forma procurada se os $v(i, j)$ forem ordenados lexicograficamente. Começando com $v(1, 1)$, subimos pela primeira coluna até $v(1, k)$, saltamos de volta para $v(2, 1)$ e subimos pela segunda coluna até onde der.

Além disso, haverá exatamente $m_k - m_{k-1}$ entradas diagonais de ordem k. Também haverá

$$\begin{array}{lll}
(m_{k-1} - m_{k-2}) - (m_k - m_{k-1}) = & 2m_{k-1} - m_k - m_{k-2} & \text{entradas diagonais de ordem } k-1 \\
 & \cdots\cdots\cdots\cdots & \\
 & 2m_2 - m_1 - m_3 & \text{entradas diagonais de ordem 2} \\
 & 2m_1 - m_2 & \text{entradas diagonais de ordem 1}
\end{array}$$

como pode ser lido diretamente da tabela. Em particular, como os números $m_1, \ldots, m_k$ são determinados de modo único por T, o número de entradas diagonais de cada ordem é determinado de modo único por T. Finalmente, a identidade

$$m_1 = (m_k - m_{k-1}) + (2m_{k-1} - m_k - m_{k-2}) + \cdots + (2m_2 - m_1 - m_3) + (2m_1 - m_2)$$

mostra que a nulidade m_1 de T é o número total de entradas diagonais de T.

10.17 Sejam $A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Não é difícil verificar que A e B são nilpotentes de índice 3, isto é, $A^3 = 0$, mas $A^2 \neq 0$, e $B^3 = 0$, mas $B^2 \neq 0$. Encontre as matrizes nilpotentes em forma canônica M_A e M_B semelhantes a A e B, respectivamente.

Como A e B são nilpotentes de índice 3, M_A e M_B devem conter um bloco de Jordan nilpotente de ordem 3 e nenhum bloco maior do que 3. Observe que pos(A) = 2 e pos(B) =3, portanto nul(A) = 5 − 2 = 3 e nul(B) = 5 − 3 = 2. Assim, M_A deve ter três blocos diagonais, que devem ser um de ordem 3 e dois de ordem 1, e M_B deve ter dois blocos diagonais, que devem ser um de ordem 3 e um de ordem 2, como segue.

$$M_A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad M_B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

10.18 Demonstre o Teorema 10.11 relativo à forma canônica de Jordan de um operador T.

Pelo teorema da decomposição primária, T é decomponível em operadores $T_1, ..., T_r$, ou seja, $T = T_1 \oplus \cdots \oplus T_r$, sendo $(t - \lambda_i)^{m_i}$ o polinômio mínimo de T_i. Assim, em particular,

$$(T_1 - \lambda_1 I)^{m_1} = 0, \quad \ldots, \quad (T_r - \lambda_r I)^{m_r} = 0$$

Denotemos $N_i = T_i - \lambda_i I$. Então, para $i = 1, ..., r$,

$$T_i = N_i + \lambda_i I, \qquad \text{com} \qquad N_i^{m_i} = 0$$

Dessa forma, T_i é a soma do operador escalar $\lambda_i I$ com um operador nilpotente N_i, que tem índice m_i, pois $(t - \lambda_i)^{m_i}$ é o polinômio mínimo de T_i.

Agora, pelo Teorema 10.10, relativo a operadores nilpotentes, podemos escolher uma base de tal forma que N_i esteja em forma canônica. Nessa base, $T_i = N_i + \lambda_i I$ é representado por uma matriz diagonal em blocos M_i cujas entradas diagonais são as matrizes J_{ij}. A soma direta J das matrizes M_i está em forma canônica de Jordan e, pelo Teorema 10.5, é uma representação matricial de T.

Finalmente, devemos mostrar que os blocos J_{ij} satisfazem as propriedades exigidas. A propriedade (i) segue do fato de que N_i tem índice m_i. A propriedade (ii) é válida porque T e J têm o mesmo polinômio característico. A propriedade (iii) é válida porque a nulidade de $N_i = T_i - \lambda_i I$ é igual à multiplicidade geométrica do autovalor λ_i. A propriedade (iv) decorre do fato de que os T_i e, portanto, os N_i, são determinados de modo único por T.

10.19 Encontre todas as possíveis formas canônicas de Jordan J de um operador linear $T: V \to V$ com polinômio característico $\Delta(t) = (t-2)^5$ e mínimo $m(t) = (t-2)^2$.

J deve ser de ordem 5, pois $\Delta(t)$ tem grau 5, e todos elementos diagonais devem ser iguais a 2, que é o único autovalor. Como o expoente de $t - 2$ em $m(t)$ é 2, J deve ter um bloco de Jordan de ordem 2 e os outros devem ser de ordem 2 ou 1. Assim, há somente duas possibilidades,

$$J = \operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, [2]\right) \qquad \text{ou} \qquad J = \operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, [2], [2], [2]\right)$$

10.20 Encontre todas as possíveis formas canônicas de Jordan de um operador linear $T: V \to V$ com polinômio característico $\Delta(t) = (t-2)^3(t-5)^2$. Em cada caso, encontre o polinômio mínimo $m(t)$.

Como $t - 2$ tem expoente 3 em $\Delta(t)$, 2, deve aparecer três vezes na diagonal. Analogamente, 5 aparece duas vezes. Assim, há seis possibilidades, como segue.

(a) $\operatorname{diag}\left(\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}, \begin{bmatrix} 5 & 1 \\ & 5 \end{bmatrix}\right)$, (b) $\operatorname{diag}\left(\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}, [5], [5]\right)$,

(c) $\operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, [2], \begin{bmatrix} 5 & 1 \\ & 5 \end{bmatrix}\right)$, (d) $\operatorname{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, [2], [5], [5]\right)$,

(e) $\operatorname{diag}\left([2], [2], [2], \begin{bmatrix} 5 & 1 \\ & 5 \end{bmatrix}\right)$, (f) $\operatorname{diag}([2], [2], [2], [5], [5])$

O expoente no polinômio mínimo $m(t)$ é igual ao tamanho do maior bloco. Assim,

(a) $m(t) = (t-2)^3(t-5)^2$, (b) $m(t) = (t-2)^3(t-5)$, (c) $m(t) = (t-2)^2(t-5)^2$,
(d) $m(t) = (t-2)^2(t-5)$, (e) $m(t) = (t-2)(t-5)^2$, (f) $m(t) = (t-2)(t-5)$

Espaço quociente e forma triangular

10.21 Seja W um subespaço de um espaço vetorial V. Mostre que as afirmações dadas são equivalentes.

(i) $u \in v + W$, (ii) $u - v \in W$, (iii) $v \in u + W$.

Suponha que $u \in v + W$. Então existe algum $w_0 \in W$ tal que $u = v + w_0$. Logo, $u - v = w_0 \in W$. Reciprocamente, suponha que $u - v \in W$. Então $u - v = w_0$ com $w_0 \in W$. Logo, $u = v + w_0 \in v + W$. Assim, são equivalentes (i) e (ii).

Também temos $u - v \in W$ se, e só se, $-(u - v) = v - u \in W$ se, e só se, $v \in u + W$. Assim, também são equivalentes (ii) e (iii).

10.22 Demonstre que as classes laterais de W em V particionam V em conjuntos dois a dois disjuntos, ou seja, prove as afirmações seguintes.

(a) Duas classes laterais $u + W$ e $v + W$ são iguais ou disjuntas.

(b) Cada $v \in V$ pertence a alguma classe, a saber, $v \in v + W$.

Além disso, $u + W = v + W$ se, e só se, $u - v \in W$, de modo que $(v + w) + W = v + W$, para cada $w \in W$.

Seja $v \in V$. Como $0 \in W$, temos $v = v + 0 \in v + W$, provando (b).
Suponha, agora, que as classes $u + W$ e $v + W$ não sejam disjuntas, digamos, que o vetor x pertença à classe $u + W$ e também à classe $v + W$. Então $u - x \in W$ e $x - v \in W$. Resta provar que $u + W = v + W$ para obter (a). Seja $u + w_0$ um elemento da classe lateral $u + W$. Como $u - x,\ x - v,\ w_0$ pertencem a W, temos

$$(u + w_0) - v = (u - x) + (x - v) + w_0 \in W$$

Assim, $u + w_0 \in v + W$, e, portanto, a classe lateral $u + W$ está contida na classe lateral $v + W$. Analogamente, $v + W$ está contido em $u + W$. Assim, $u + W = v + W$.
Finalmente, $u + W = v + W$ se, e só se, $u \in v + W$ e, pelo Problema 10.21, isso equivale a $u - v \in W$. A última afirmação está demonstrada.

10.23 Seja W o espaço solução da equação homogênea $2x + 3y + 4z = 0$. Descreva as classes laterais de W em $\mathbf{R}^3$.

W é um plano pela origem $O = (0, 0, 0)$ e as classes laterais de W são os planos paralelos a W. Equivalentemente, as classes laterais de W são os conjuntos solução da família de equações

$$2x + 3y + 4z = k, \qquad k \in \mathbf{R}$$

De fato, a classe $v + W$, com $v = (a, b, c)$, é o conjunto solução da equação linear

$$2x + 3y + 4z = 2a + 3b + 4c \qquad \text{ou} \qquad 2(x - a) + 3(y - b) + 4(z - c) = 0$$

10.24 Seja W um subespaço de um espaço vetorial V. Mostre que as operações dadas no Teorema 10.15 estão bem definidas, ou seja, mostre que se $u + W = u' + W$ e $v + W = v' + W$, então

(a) $(u + v) + W = (u' + v') + W$ e (b) $ru + W = ru' + W$, com $r \in K$

(a) Como $u + W = u' + W$ e $v + W = v' + W$, tanto $u - u'$ quanto $v - v'$ pertencem a W. Segue que $(u + v) - (u' + v') = (u - u') + (v - v') \in W$ e, portanto, $(u + v) + W = (u' + v') + W$.

(b) Também, como $u - u' \in W$ implica $r(u - u') \in W$, então $ru - ru' = r(u - u') \in W$ e, dessa forma, $ru + W = ru' + W$.

10.25 Sejam V um espaço vetorial e W um subespaço de V. Mostre que a aplicação natural $\eta\colon V \to V/W$, definida por $\eta(v) = v + W$, é linear.

Dados $u, v \in V$ e $k \in K$ quaisquer, temos

$$n(u + v) = u + v + W = u + W + v + W = \eta(u) + \eta(v)$$

e

$$\eta(kv) = kv + W = k(v + W) = k\eta(v)$$

Assim, η é linear.

10.26 Seja $\{w_1, ..., w_r\}$ uma base de um subespaço W de um espaço vetorial V. Suponha que o conjunto de classes laterais $\{\bar{v}_1, \ldots, \bar{v}_s\}$, com $\bar{v}_j = v_j + W$, seja uma base do espaço quociente V/W. Mostre que o conjunto $B = \{v_1, \ldots, v_s, w_1, \ldots, w_r\}$ de vetores é uma base de V. Assim, $\dim V = \dim W + \dim(V/W)$.

Seja $u \in V$. Como $\{\bar{v}_j\}$ é uma base de V/W, temos

$$\bar{u} = u + W = a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_s\bar{v}_s$$

Logo, $u = a_1v_1 + \cdots + a_sv_s + w$, com $w \in W$. Como $\{w_i\}$ é uma base de W,

$$u = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_rw_r$$

Desse modo, B gera V.

Mostremos que B é linearmente independente. Suponha que

$$c_1 v_1 + \cdots + c_s v_s + d_1 w_1 + \cdots + d_r w_r = 0 \tag{1}$$

Então

$$c_1 \bar{v}_1 + \cdots + c_s \bar{v}_s = \bar{0} = W$$

Como $\{\bar{v}_j\}$ é independente, os c são todos nulos. Substituindo em (1), obtemos $d_1 w_1 + \cdots + d_r w_r = 0$. Como $\{w_i\}$ é independente, os d são todos nulos. Assim, B é linearmente independente e, portanto, uma base de V.

10.27 Demonstre o Teorema 10.16. Seja W um subespaço invariante por um operador linear $T{:}V \to V$. Então T induz um operador linear $\bar{T}$ de V/W definido por $\bar{T}(v + W) = T(v) + W$. Além disso, se T é uma raiz de algum polinômio, então $\bar{T}$ também é. Assim, o polinômio mínimo de $\bar{T}$ divide o polinômio mínimo de T.

Começamos mostrando que $\bar{T}$ está bem definido, isto é, se $u + W = v + W$, então $\bar{T}(u + W) = \bar{T}(v + W)$. Se $u + W = v + W$, então $u - v \in W$ e, como W é T-invariante, $T(u - v) = T(u) - T(v) \in W$. Dessa forma,

$$\bar{T}(u + W) = T(u) + W = T(v) + W = \bar{T}(v + W)$$

como queríamos mostrar.

Em seguida, mostramos que $\bar{T}$ é linear. Temos

$$\begin{aligned}\bar{T}((u + W) + (v + W)) &= \bar{T}(u + v + W) = T(u + v) + W = T(u) + T(v) + W \\ &= T(u) + W + T(v) + W = \bar{T}(u + W) + \bar{T}(v + W)\end{aligned}$$

Além disso,

$$\bar{T}(k(u + W)) = \bar{T}(ru + W) = T(ru) + W = kT(u) + W = k(T(u) + W) = k\bar{T}(u + W)$$

Assim, $\bar{T}$ é linear.

Finalmente, seja $u + W$ uma classe lateral qualquer de V/W. Temos

$$\overline{T^2}(u + W) = T^2(u) + W = T(T(u)) + W = \bar{T}(T(u) + W) = \bar{T}(\bar{T}(u + W)) = \bar{T}^2(u + W)$$

Logo, $\overline{T^2} = \bar{T}^2$. Analogamente, $\overline{T^n} = \bar{T}^n$, para cada n. Assim, dado um polinômio qualquer

$$f(t) = a_n t^n + \cdots + a_0 = \sum a_i t^i$$

$$\begin{aligned}\overline{f(T)}(u + W) &= f(T)(u) + W = \sum a_i T^i(u) + W = \sum a_i (T^i(u) + W) \\ &= \sum a_i \overline{T^i}(u + W) = \sum a_i \bar{T}^i(u + W) = (\sum a_i \bar{T}^i)(u + W) = f(\bar{T})(u + W)\end{aligned}$$

e, portanto, $\overline{f(T)} = f(\bar{T})$. Dessa forma, se T é uma raiz de f(t), então $\overline{f(T)} = \bar{0} = W = f(\bar{T})$; ou seja, $\bar{T}$ também é uma raiz de $f(t)$, provando o teorema.

10.28 Demonstre o Teorema 10.1. Seja $T{:}V \to V$ um operador linear cujo polinômio característico fatora em polinômios lineares. Então existe uma base de V na qual T é representado por uma matriz triangular.

A demonstração é por indução na dimensão de V. Se dim $V = 1$, toda representação matricial de T é uma matriz 1×1, que é triangular.

Agora suponha que dim $V = n > 1$ e que o teorema seja válido para espaços de dimensões menores do que n. Como o polinômio característico de T fatora em polinômios lineares, T tem, pelo menos, um autovalor e, portanto, pelo menos um autovetor v não nulo, digamos que $T(v) = a_{11} v$. Seja W o subespaço unidimensional gerado por v e denotemos $\bar{V} = V/W$. Então (Problema 10.26) dim $\bar{V} = \dim V - \dim W = n - 1$. Observe, também, que W é invariante por T. Pelo Teorema 10.16, T induz um operador linear $\bar{T}$ de $\bar{V}$ cujo polinômio mínimo divide o polinômio mínimo de T. Como o polinômio característico de T é um produto de polinômios lineares, o mesmo ocorre com seu polinômio mínimo e, portanto, com os polinômios mínimo e característico de $\bar{T}$. Assim, $\bar{V}$ e $\bar{T}$ satisfazem as hipóteses do teorema. Logo, por indução, existe uma base $\{\bar{v}_2, \ldots, \bar{v}_n\}$ de $\bar{V}$ tal que

$$\begin{aligned}&\bar{T}(\bar{v}_2) = a_{22}\bar{v}_2 \\ &\bar{T}(\bar{v}_3) = a_{32}\bar{v}_2 + a_{33}\bar{v}_3 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\bar{T}(\bar{v}_n) = a_{n2}\bar{v}_n + a_{n3}\bar{v}_3 + \cdots + a_{nn}\bar{v}_n\end{aligned}$$

Sejam, agora, $v_2, ..., v_n$ elementos de V que pertencem às classes laterais $\bar{v}_2, \ldots, \bar{v}_n$, respectivamente, Então $\{v_2, ..., v_n\}$ é uma base de V (Problema 10.26). Como $\bar{T}(\bar{v}_2) = a_{22}\bar{v}_2$, temos

$$\bar{T}(\bar{v}_2) - a_{22}\bar{v}_{22} = 0, \qquad \text{e, portanto,} \qquad T(v_2) - a_{22}v_2 \in W$$

Mas W é gerado por v, portanto, $T(v_2) - a_{22}v_2$ é um múltiplo de v, digamos,

$$T(v_2) - a_{22}v_2 = a_{21}v, \qquad \text{e, portanto,} \qquad T(v_2) = a_{21}v + a_{22}v_2$$

Analogamente para $i = 3, ..., n$,

$$T(v_i) - a_{i2}v_2 - a_{i3}v_3 - \cdots - a_{ii}v_i \in W, \qquad \text{e, portanto,} \qquad T(v_i) = a_{i1}v + a_{i2}v_2 + \cdots + a_{ii}v_i$$

Assim,

$$\begin{aligned} T(v) &= a_{11}v \\ T(v_2) &= a_{21}v + a_{22}v_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ T(v_n) &= a_{n1}v + a_{n2}v_2 + \cdots + a_{nn}v_n \end{aligned}$$

e, dessa forma, nessa base, a matriz de T é triangular.

Subespaços cíclicos, forma canônica racional

10.29 Demonstre o Teorema 10.12. Sejam $Z(v, T)$ um subespaço T-cíclico, T_v a restrição de T a $Z(v, T)$, e $m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_0$ o T-anulador de v. Valem as afirmações seguintes.

(i) O conjunto $\{v, T(v), \ldots, T^{k-1}(v)\}$ é uma base de $Z(v, T)$, portanto, dim $Z(v, T) = k$.

(ii) O polinômio mínimo de T_v é $m_v(t)$.

(iii) A representação matricial de T_v na base de (i) é a matriz companheira $C = C(m_v)$ de $m_v(t)$ [que tem entradas iguais a 1 na subdiagonal, o simétrico dos coeficientes $a_0, a_1, ..., a_{k-1}$ de $m_v(t)$ na última coluna, e entradas nulas no resto].

(i) Por definição de $m_v(t)$, o vetor $T^k(v)$ é o primeiro da sequência $v, T(v), T^2(v), ...$ que é uma combinação linear dos vetores que o precedem na sequência. Assim, o conjunto $B = \{v, T(v), T^2(v), ..., T^{k-1}(v)\}$ é linearmente independente. Resta mostrar que $Z(v, T) = \text{ger}(B)$, o espaço gerado por B. Pelo que acabamos dever, $T^k(v) \in \text{ger}(B)$. Usamos indução para provar que $T^n(v) \in \text{ger}(B)$, para todo n. Supomos que $n > k$ e que $T^{n-1}(v) \in \text{ger}(B)$, ou seja, que $T^{n-1}(v)$ é uma combinação linear de $v, T(v), T^2(v), ..., T^{k-1}(v)$. Então $T^n(v) = T(T^{n-1}(v))$ é uma combinação linear de $T(v), T^2(v), ..., T^k(v)$. Como $T^k(v) \in \text{ger}(B)$, segue que $T^n(v) \in \text{ger}(B)$, para cada n. Consequentemente, $f(T)(v) \in \text{ger}(B)$ para qualquer polinômio $f(t)$. Assim, $Z(v, T) = \text{ger}(B)$ e, portanto, B é uma base, como queríamos mostrar.

(ii) Seja $m(t) = t^s + b_{s-1}t^{s-1} + \cdots + b_0$ o polinômio mínimo de T_v. Como $v \in Z(v, T)$, temos

$$0 = m(T_v)(v) = m(T)(v) = T^s(v) + b_{s-1}T^{s-1}(v) + \cdots + b_0 v$$

Assim, $T^s(v)$ é uma combinação linear de $v, \; T(v), \ldots, T^{s-1}(v)$, de modo que $k \leq s$. Contudo, $m_v(T) = 0$ e, portanto, $m_v(T_v) = 0$. Então $m(t)$ divide $m_v(t)$, de modo que $s \leq k$. Decorre que $k = s$ e, portanto, $m_v(t) = m(t)$.

(iii)

$$\begin{aligned} T_v(v) &= T(v) \\ T_v(T(v)) &= T^2(v) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ T_v(T^{k-2}(v)) &= T^{k-1}(v) \\ T_v(T^{k-1}(v)) &= T^k(v) = -a_0 v - a_1 T(v) - a_2 T^2(v) - \cdots - a_{k-1}T^{k-1}(v) \end{aligned}$$

Por definição, a matriz de T_v nessa base é a transposta da matriz de coeficientes desse sistema de equações lineares, ou seja, é C, como queríamos mostrar.

10.30 Seja $T:V \to V$ um operador linear. Seja W um subespaço T-invariante de V e $\bar{T}$ o operador induzido no quociente V/W. Prove as afirmações dadas.

(a) O T-anulador de $v \in V$ divide o polinômio mínimo de T.

(b) O $\bar{T}$-anulador de $\bar{v} \in V/W$ divide o polinômio mínimo de T.

(a) O T-anulador de $v \in V$ é o polinômio mínimo da restrição de T a $Z(v, T)$, portanto, pelo Problema 10.6, divide o polinômio mínimo de T.

(b) O $\bar{T}$-anulador de $\bar{v} \in V/W$ divide o polinômio mínimo de $\bar{T}$, que divide o polinômio mínimo de T pelo Teorema 10.16.

OBSERVAÇÃO Se o polinômio mínimo de T é $f(t)^n$, com $f(t)$ um polinômio mônico irredutível, então o T-anulador de $v \in V$ e o $\bar{T}$-anulador de $\bar{v} \in V/W$ são da forma $f(t)^m$, com $m \leq n$.

10.31 Demonstre o Lema 10.13. Seja $T:V \to V$ um operador linear cujo polinômio mínimo é $f(t)^n$, em que $f(t)$ é um polinômio mônico irredutível. Então V é a soma direta de subespaços T-cíclicos $Z_i = Z(v_i, T)$, em que $i = 1, ..., r$, com T-anuladores correspondentes

$$f(t)^{n_1}, f(t)^{n_2}, \ldots, f(t)^{n_r}, \qquad n = n_1 \geq n_2 \geq \cdots \geq n_r$$

Qualquer outra decomposição de V numa soma direta de subespaços T-cíclicos tem o mesmo número de componentes e o mesmo conjunto de T-anuladores.

A prova é por indução na dimensão de V. Se dim $V = 1$, então V é T-cíclico e vale o lema. Agora suponha que dim $V > 1$ e que o lema seja válido para todos os espaços vetoriais de dimensão inferior à de V.

Como o polinômio mínimo de T é $f(t)^n$, existe algum $v_1 \in V$ tal que $f(T)^{n-1}(v_1) \neq 0$; de modo que o T-anulador de v_1 é $f(t)^n$. Seja $Z_1 = Z(v_1, T)$; sabemos que Z_1 é T-invariante. Denotemos $\bar{V} = V/Z_1$ e seja $\bar{T}$ o operador linear de $\bar{V}$ induzido por T. Pelo Teorema 10.16, o polinômio mínimo de $\bar{T}$ divide $f(t)^n$. Logo, a hipótese vale para $\bar{V}$ e $\bar{T}$ e, consequentemente, por indução, $\bar{V}$ é a soma direta de subespaços $\bar{T}$-cíclicos, digamos,

$$\bar{V} = Z(\bar{v}_2, \bar{T}) \oplus \cdots \oplus Z(\bar{v}_r, \bar{T})$$

em que os $\bar{T}$-anuladores correspondentes são $f(t)^{n_2}, \ldots, f(t)^{n_r}$, $n \geq n_2 \geq \cdots \geq n_r$.

Afirmamos que existe um vetor v_2 na classe lateral $\bar{v}_2$ cujo T-anulador é $f(t)^{n_2}$, o $\bar{T}$-anulador da classe $\bar{v}_2$. Seja w um vetor qualquer de $\bar{v}_2$. Então $f(T)^{n_2}(w) \in Z_1$. Logo, existe algum polinômio $g(t)$ para o qual

$$f(T)^{n_2}(w) = g(T)(v_1) \tag{1}$$

Como $f(t)^n$ é o polinômio mínimo de T, (1) garante que

$$0 = f(T)^n(w) = f(T)^{n-n_2} g(T)(v_1)$$

No entanto, $f(t)^n$ é o T-anulador de v_1, portanto, $f(t)^n$ divide $f(t)^{n-n_2}g(t)$, e, dessa forma, $g(t) = f(t)^{n_2}h(t)$, para algum polinômio $h(t)$. Denotamos

$$v_2 = w - h(T)(v_1)$$

Como $w - v_2 = h(T)(v_1) \in Z_1$, v_2, também pertence à classe lateral $\bar{v}_2$. Assim, o T-anulador de v_2 é um múltiplo do $\bar{T}$-anulador de $\bar{v}_2$. Por outro lado, por (1),

$$f(T)^{n_2}(v_2) = f(T)^{n_s}(w - h(T)(v_1)) = f(T)^{n_2}(w) - g(T)(v_1) = 0$$

Consequentemente, o T-anulador de v_2 é v_2 is $f(t)^{n_2}$, como afirmamos.

Analogamente, existem $v_3, \ldots, v_r \in V$ tais que $v_i \in \overline{v_i}$ e tais que o T-anulador de v_i é $f(t)^{n_i}$, o $\bar{T}$-anulador de $\overline{v_i}$. Definimos

$$Z_2 = Z(v_2, T), \qquad \ldots, \qquad Z_r = Z(v_r, T)$$

Seja d o grau de $f(t)$, de modo que o grau de $f(t)^{n_i}$ é dn_i. Então, por ser $f(t)^{n_i}$ o T-anulador de v_i e também o $\bar{T}$-anulador de $\overline{v_i}$, sabemos que

$$\{v_i, T(v_i), \ldots, T^{dn_i-1}(v_i)\} \qquad \text{e} \qquad \{\bar{v}_i, \bar{T}(\overline{v_i}), \ldots, \bar{T}^{dn_i-1}(\overline{v_i})\}$$

são bases de $Z(v_i, T)$ e $Z(\overline{v_i}, \bar{T})$, respectivamente, para $i = 2, ..., r$. No entanto, $\bar{V} = Z(\overline{v_2}, \bar{T}) \oplus \cdots \oplus Z(\overline{v_r}, \bar{T})$, de modo que

$$\{\bar{v}_2, \ldots, \bar{T}^{dn_2-1}(\bar{v}_2), \ldots, \bar{v}_r, \ldots, \bar{T}^{dn_r-1}(\bar{v}_r)\}$$

é uma base de $\bar{V}$. Portanto, pelo Problema 10.26 e a relação $\bar{T}^i(\bar{v}) = \overline{T^i(v)}$ (ver Problema 10.27),

$$\{v_1, \ldots, T^{dn_1-1}(v_1), v_2, \ldots, T^{en_2-1}(v_2), \ldots, v_r, \ldots, T^{dn_r-1}(v_r)\}$$

é uma base de V. Assim, pelo Teorema 10.4, $V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$, como queríamos mostrar.

Resta provar que os expoentes n_1, ..., n_r são determinados de modo único por T. Como d é o grau de $f(t)$,

$$\dim V = d(n_1 + \cdots + n_r) \quad \text{e} \quad \dim Z_i = dn_i, \quad i = 1, \ldots, r$$

Também, se s é um inteiro positivo qualquer, então (Problema 10.59) $f(T)^s(Z_i)$ é um subespaço cíclico gerado por $f(T)^s(v_i)$, que tem dimensão $d(n_i - s)$ se $n_i > s$ e dimensão 0 se $n_i \leq s$.

Agora, qualquer vetor $v \in V$ pode ser escrito de modo único na forma $v = w_1 + \cdots + w_r$, onde $w_i \in Z_i$. Logo, qualquer vetor de $f(T)^s(V)$ pode ser escrito, de modo único, na forma

$$f(T)^s(v) = f(T)^s(w_1) + \cdots + f(T)^s(w_r)$$

com $f(T)^s(w_i) \in f(T)^s(Z_i)$. Seja t o inteiro (que depende de s) para o qual

$$n_1 > s, \quad \ldots, \quad n_t > s, \quad n_{t+1} \geq s$$

Então

$$f(T)^s(V) = f(T)^s(Z_1) \oplus \cdots \oplus f(T)^s(Z_t)$$

e, portanto,

$$\dim[f(T)^s(V)] = d[(n_1 - s) + \cdots + (n_t - s)] \tag{2}$$

Os números à esquerda de (2) são determinado de modo único por T. Tomando $s = n - 1$, (2) determina o número de n_i que são iguais a n. Em seguida, tomando $s = n - 2$, (2) determina o número de n_i (se existirem) que são iguais a $n - 1$. Repetimos esse processo até tomar $s = 0$ e determinar o número de n_i que são iguais a 1. Assim, os n_i são determinados, de modo único por T e V, terminando a demonstração do lema.

10.32 Sejam V um espaço vetorial de dimensão sete sobre $\mathbf{R}$ e $T: V \to V$ um operador linear de polinômio mínimo $m(t) = (t^2 - 2t + 5)(t - 3)^3$. Encontre todas as formas canônicas racionais M possíveis de T.

Como $\dim V = 7$, só há duas possibilidades para o polinômio característico de T, $\Delta_1(t) = (t^2 - 2t + 5)^2 (t - 3)^3$ ou $\Delta_1(t) = (t^2 - 2t + 5)(t - 3)^5$. Além disso, a soma das ordens das matrizes companheiras deve totalizar 7. Também uma matriz companheira deve ser $C(t^2 - 2t + 5)$ e uma outra deve ser $C((t - 3)^3) = C(t^3 - 9t^2 + 27t - 27)$. Assim, M deve ser uma das três matrizes diagonais em bloco a seguir.

(a) $\operatorname{diag}\left(\begin{bmatrix} 0 & -5 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & -5 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 27 \\ 1 & 0 & -27 \\ 0 & 1 & 9 \end{bmatrix} \right),$

(b) $\operatorname{diag}\left(\begin{bmatrix} 0 & -5 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 27 \\ 1 & 0 & -27 \\ 0 & 1 & 9 \end{bmatrix}, \begin{bmatrix} 0 & -9 \\ 1 & 6 \end{bmatrix} \right),$

(c) $\operatorname{diag}\left(\begin{bmatrix} 0 & -5 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 27 \\ 1 & 0 & -27 \\ 0 & 1 & 9 \end{bmatrix}, [3], [3] \right)$

Projeções

10.33 Suponha que $V = W_1 \oplus \cdots \oplus W_r$. A *projeção* de V sobre o subespaço W_k é a aplicação $E: V \to V$ definida por $E(v) = w_k$, se $v = w_1 + \cdots + w_r$, $w_i \in W_i$. Mostre que (a) E é um operador linear, (b) $E^2 = E$.

(a) Como a soma $v = w_1 + \cdots + w_r$, $w_i \in W$ é determinada, de modo único, por v, a aplicação E está bem definida. Suponha que, dado $u \in V$, tenhamos $u = w'_1 + \cdots + w'_r$, $w'_i \in W_i$. Então

$$v + u = (w_1 + w'_1) + \cdots + (w_r + w'_r) \quad \text{e} \quad kv = kw_1 + \cdots + kw_r, \quad kw_i, w_i + w'_i \in W_i$$

são as somas únicas correspondentes a $v + u$ e kv. Então,

$$E(v + u) = w_k + w'_k = E(v) + E(u) \quad \text{e} \quad E(kv) = kw_k + kE(v)$$

e, portanto, E é linear.

(b) Temos que

$$w_k = 0 + \cdots + 0 + w_k + 0 + \cdots + 0$$

é a única soma correspondente a $w_k \in W_k$ e, portanto, $E(w_k) = w_k$. Então, para cada $v \in V$,

$$E^2(v) = E(E(v)) = E(w_k) = w_k = E(v)$$

Assim, $E^2 = E$, como queríamos mostrar.

10.34 Suponha que $E:V \to V$ seja um operador linear tal que $E^2 = E$. Mostre que (a) $E(u) = u$, para cada $u \in \text{Im } E$ (isto é, $V = \text{Im } E \oplus \text{Nuc } E$ restrição de E à imagem de E é a transformação identidade); (b) V é a soma direta da imagem e do núcleo de E, isto é, $V = \text{Im } E \oplus \text{Nuc } E$; (c) E é a projeção de V sobre Im E, a imagem de E. Assim, pelo problema precedente, um operador linear $T:V \to V$ é uma projeção se, e só se, $T^2 = T$. Essa caracterização de projeção é usada frequentemente como definição.

(a) Se $u \in \text{Im } E$, então existe algum $v \in V$ tal que $E(v) = u$. Segue que

$$E(u) = E(E(v)) = E^2(v) = E(v) = u$$

(b) Seja $v \in V$. Podemos escrever v no formato $v = E(v) + v - E(v)$. Como $E(v) \in \text{Im } E$ e, como

$$E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

temos $v - E(v) \in \text{Nuc } E$. Assim, $V = \text{Im } E + \text{Nuc } E$.
Agora, suponha que $w \in \text{Im } E \cap \text{Nuc} E$. Por (a), temos $E(w) = w$, porque $w \in \text{Nuc } E$. Por outro lado, $E(w) = 0$, já que $w \in \text{Nuc } E$. Assim, $w = 0$ e, portanto, $\text{Im } E \cap \text{Nuc} E = \{0\}$. Essas duas condições implicam que V é a soma direta da imagem e do núcleo de E.

(c) Seja $v \in V$ e suponha que $v = u + w$, com $u \in \text{Im } E$ e $w \in \text{Nuc } E$. Observe que $E(u) = u$, por (a), e que $E(w) = 0$, porque $w \in \text{Nuc } E$. Logo,

$$E(v) = E(u + w) = E(u) + E(w) = u + 0 = u$$

Assim, E é a projeção de V sobre a imagem de E.

10.35 Suponha que $V = U \oplus W$ e seja $T:V \to V$ um operador linear. Mostre que U e W são T-invariantes se, e só se, $TE = ET$, onde E é projeção de V sobre U.

Observe que $E(v) \in U$, para cada $v \in V$, e que (i) $E(v) = v$ se, e só se, $v \in U$, (ii) $E(v) = 0$ se, e só se, $v \in W$. Suponha que $ET = TE$. Seja $u \in U$. Como $E(u) = u$,

$$T(u) = T(E(u)) = (TE)(u) = (ET)(u) = E(T(u)) \in U$$

Logo, U é T-invariante. Seja, agora, $w \in W$. Como $E(w) = 0$,

$$E(T(w)) = (ET)(w) = (TE)(w) = T(E(w)) = T(0) = 0, \text{ e, portanto, } T(w) \in W$$

Assim, W também é T-invariante.
Reciprocamente, sejam U e W subespaços T-invariantes. Seja $v \in V$ e suponha que $v = u + w$, com $u \in T$ e $w \in W$. Então, $T(u) \in U$ e $T(w) \in W$. Logo, $E(T(u)) = T(u)$ e $E(T(w)) = 0$. Assim,

$$(ET)(v) = (ET)(u + w) = (ET)(u) + (ET)(w) = E(T(u)) + E(T(w)) = T(u)$$

e

$$(TE)(v) = (TE)(u + w) = T(E(u + w)) = T(u)$$

Ou seja, $(ET)(v) = (TE)(v)$, para cada $v \in V$; assim, $ET = TE$, como queríamos mostrar.

Problemas Complementares

Subespaços invariantes

10.36 Sejam W invariante por $T:V \to V$ e $f(t)$ um polinômio. Mostre que W é invariante por $f(T)$.

10.37 Mostre que cada subespaço de V é invariante por I e 0, os operadores identidade e nulo, respectivamente.

10.38 Seja W invariante por $T_1: V \to V$ e $T_2: V \to V$. Prove que W também é invariante por $T_1 + T_2$ e T_1T_2.

10.39 Seja $T:V \to V$ um operador linear. Prove que cada autoespaço E_λ de T é T-invariante.

10.40 Seja V um espaço vetorial de dimensão ímpar (maior do que 1) sobre o corpo dos reais $\mathbf{R}$. Mostre que qualquer operador linear de V tem algum subespaço invariante além de V e $\{0\}$.

10.41 Encontre todos os subespaços invariantes por $A = \begin{bmatrix} 2 & -4 \\ 5 & -2 \end{bmatrix}$ vista como um operador linear de (a) $\mathbf{R}^2$, (b) $\mathbf{C}^2$.

10.42 Suponha que dim $V = n$. Mostre que $T:V \to V$ tem uma representação matricial triangular se, e só se, existem subespaços T-invariantes $W_1 \subset W_2 \subset \cdots \subset W_n = V$ tais que dim $W_k = k$, com $k = 1, ..., n$.

Somas diretas invariantes

10.43 Dizemos que os subespaços $W_1, ..., W_r$ são *independentes* se $w_1 + \cdots + w_r = 0$, $w_i \in W_i$, implica que cada $w_i = 0$. Mostre que ger$(W_i) = W_1 \oplus \cdots \oplus W_r$ se, e só se, os W_i são independentes. [Aqui, ger(W_i) denota o espaço gerado pelos W_i.]

10.44 Mostre que $V = W_1 \oplus \cdots \oplus W_r$ se, e só se, (i) V = ger$(W_1, ..., W_r)$ e (ii) para cada $k = 1, ..., r$, vale $W_k \cap \text{ger}\,(W_1, \ldots, W_{k-1}, W_{k+1}, \ldots, W_r) = \{0\}$.

10.45 Mostre que $\text{ger}\,(W_i) = W_1 \oplus \cdots \oplus W_r$ se, e só se, $\dim[\text{ger}\,(W_i)] = \dim W_1 + \cdots + \dim W_r$.

10.46 Suponha que o polinômio característico de $T:V \to V$ seja $\Delta(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_r(t)^{n_r}$, em que cada $f_i(t)$ é um polinômio mônico irredutível. Seja $V = W_1 \oplus \cdots \oplus W_r$ a decomposição primária de V em subespaços T-invariantes. Mostre que $f_i(t)^{n_i}$ é o polinômio característico da restrição de T a W_i.

Operadores nilpotentes

10.47 Sejam T_1 e T_2 operadores nilpotentes que comutam, isto é, $T_1T_2 = T_2T_1$. Mostre que $T_1 + T_2$ e T_1T_2 também são nilpotentes.

10.48 Seja A uma matriz supertriangular, isto é, todas as entradas na diagonal e abaixo da diagonal são nulas. Mostre que A é nilpotente.

10.49 Seja V o espaço vetorial dos polinômios de grau $\leq n$. Mostre que o operador derivada de V é nilpotente de índice $n + 1$.

10.50 Mostre que qualquer bloco de Jordan nilpotente N é semelhante à sua transposta N^T (que tem entradas iguais a 1 na subdiagonal e demais entradas nulas).

10.51 Mostre que duas matrizes nilpotentes de ordem 3 são semelhantes se, e só se, tiverem o mesmo índice de nilpotência. Mostre com um exemplo que a afirmação não é verdadeira para matrizes nilpotentes de ordem 4.

Forma canônica de Jordan

10.52 Encontre todas as possíveis formas canônicas de Jordan das matrizes cujos polinômios característico $\Delta(t)$ e mínimo $m(t)$ estão dados.
(a) $\Delta(t) = (t-2)^4(t-3)^2$, $m(t) = (t-2)^2(t-3)^2$,
(b) $\Delta(t) = (t-7)^5$, $m(t) = (t-7)^2$, (c) $\Delta(t) = (t-2)^7$, $m(t) = (t-2)^3$

10.53 Mostre que cada matriz complexa é semelhante à sua transposta. (*Sugestão*: use a forma canônica de Jordan.)

10.54 Mostre que são semelhantes todas as matrizes $n \times n$ complexas A tais que $A^n = I$, mas $A_k \neq I$, com $k < n$.

10.55 Seja A uma matriz complexa com todos os autovalores reais. Mostre que A é semelhante a uma matriz com todas entradas reais.

Subespaços cíclicos

10.56 Seja $T:V \to V$ um operador linear de V. Prove que $Z(v, T)$ é a interseção de todos os subespaços T-invariantes de V que contêm v.

10.57 Sejam $f(t)$ e $g(t)$ os T-anuladores de u e v, respectivamente. Mostre que se $f(t)$ e $g(t)$ são primos entre si, então $f(t)g(t)$ é o T-anulador de $u + v$.

10.58 Prove que $Z(u, T) = Z(v, T)$ se, e só se, $g(T)(u) = v$, para algum polinômio $g(t)$ primo com o polinômio T-anulador de u.

10.59 Seja $W = Z(v, T)$ e suponha que o T-anulador de v seja $f(t)^n$, em que $f(t)$ é um polinômio mônico irredutível de grau d. Mostre que $f(T)^s(W)$ é o subespaço cíclico gerado por $f(T)^s(v)$ e que tem dimensão $d(n - s)$ se $n > s$ e dimensão 0 se $n \leq s$.

Forma canônica racional

10.60 Encontre todas as formas canônicas racionais possíveis para matrizes 6×6 sobre $\mathbf{R}$ de polinômio mínimo dado.
(a) $m(t) = (t^2 - 2t + 3)(t + 1)^2$, (b) $m(t) = (t - 2)^3$.

10.61 Seja A uma matriz 4×4 de polinômio mínimo $m(t) = (t^2 + 1)(t^2 - 3)$. Encontre a forma canônica racional de A se A é uma matriz sobre o corpo dos (a) racionais $\mathbf{Q}$, (b) reais $\mathbf{R}$, (c) complexos $\mathbf{C}$.

10.62 Encontre a forma canônica racional de um bloco de Jordan de ordem 4 com λ na diagonal.

10.63 Prove que o polinômio característico de um operador $T:V \to V$ é um produto de seus divisores elementares.

10.64 Prove que são semelhantes duas matrizes 3×3 com os mesmos polinômios mínimo e característico.

10.65 Denote por $C(f(t))$ a matriz companheira de um polinômio $f(t)$ arbitrário. Mostre que $f(t)$ é o polinômio característico de $C(f(t))$.

Projeções

10.66 Suponha que $V = W_1 \oplus \cdots \oplus W_r$. Seja E_i a projeção de V sobre W_i. Prove que (i) $E_iE_j = 0$, $i \neq j$; (ii) $I = E_1 + \cdots + E_r$.

10.67 Sejam $E_1, \ldots, E_r$ operadores lineares de V tais que

(i) $E_i^2 = E_i$ (ou seja, cada E_i é uma projeção); (ii) $E_iE_j = 0$, $i \neq j$; (iii) $I = E_1 + \cdots + E_r$

Prove que $V = \text{Im}\, E_1 \oplus \cdots \oplus \text{Im}\, E_r$.

10.68 Seja $E: V \to V$ uma projeção (ou seja, $E^2 = E$). Prove que E tem uma representação matricial da forma $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, em que r é o posto de E e I_r é a matriz identidade de ordem r.

10.69 Prove que duas projeções de mesmo posto são semelhantes. (*Sugestão*: use o resultado do Problema 10.68.)

10.70 Seja $E: V \to V$ uma projeção. Prove

(i) $I - E$ é uma projeção e $V = \text{Im}\, E \oplus \text{Im}\, (I - E)$, (ii) $I + E$ é invertível (se $1 + 1 \neq 0$).

Espaços quociente

10.71 Seja W um subespaço de V. Suponha que um conjunto $\{v_1 + W, \; v_2 + W, \; \ldots, \; v_n + W\}$ de classes laterais de V/W seja linearmente independente. Mostre que o conjunto $\{v_1, v_2, \ldots, v_n\}$ de vetores de V também é linearmente independente.

10.72 Seja W um subespaço de V. Suponha que um conjunto $\{u_1, u_2, \ldots, u_n\}$ de vetores de V seja linearmente independente e que $\text{ger}(u_i) \cap W = \{0\}$. Mostre que o conjunto $\{u_1 + W, \ \ldots, \ u_n + W\}$ de classes laterais de V/W também é linearmente independente.

10.73 Suponha que $V = U \oplus W$ e que $\{u_1, ..., u_n\}$ seja uma base de U. Mostre que $\{u_1 + W, ..., u_n + W\}$ é uma base do espaço quociente V/W. (Observe que não se impõe restrição sobre a dimensionalidade de V ou W.)

10.74 Seja W o espaço solução da equação linear

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0, \qquad a_i \in K$$

e seja $v = (b_1, b_2, \ldots, b_n) \in K^n$. Prove que a classe lateral $v + W$ de W em K^n é o conjunto solução da equação linear

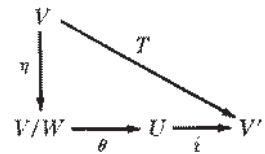
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \qquad \text{com} \qquad b = a_1b_1 + \cdots + a_nb_n$$

10.75 Seja V o espaço vetorial de polinômios sobre **R** e W o subespaço dos polinômios divisíveis por t^4 (ou seja, da forma $a_0t^4 + a_1t^5 + \cdots + a_{n-4}t^n$). Mostre que o espaço quociente V/W tem dimensão 4.

10.76 Sejam U e W subespaços de V tais que $W \subset U \subset V$. Observe que cada classe lateral $u + W$ de W em U também pode ser vista como uma classe lateral de W em V, porque $u \in U$ implica $u \in V$. Assim, U/W é um subconjunto de V/W. Prove que (i) U/W é um subespaço de V/W, (ii) $\dim(V/W) - \dim(U/W) = \dim(V/U)$.

10.77 Sejam U e W subespaços de V. Mostre que as classes laterais de $U \cap W$ em V podem ser obtidas intersectando cada classe lateral de U em V por cada classe lateral de W em V, ou seja,

$$V/(U \cap W) = \{(v + U) \cap (v' + W) : v, v' \in V\}$$

10.78 Seja $T: V \to V'$ um operador linear de núcleo W e imagem U. Mostre que o espaço quociente V/W é isomorfo a U pela aplicação $\theta: V/W \to U$ definida por $\theta(v + W) = T(v)$. Além disso, mostre que $T = i \circ \theta \circ \eta$, onde $\eta: V \to V/W$ é a aplicação natural de V em V/W (ou seja, $\eta(v) = v + W$) e $i: U \hookrightarrow V'$ é a aplicação de inclusão $i(u) = u$ (ou seja, $i(u) = u$). (Ver diagrama.)

Respostas dos Problemas Complementares

10.41 (a) $\mathbf{R}^2$ e $\{0\}$, (b) $\mathbf{C}^2$, $\{0\}$, $W_1 = \text{ger}(2, 1 - 2i)$, $W_2 = \text{ger}(2, 1 + 2i)$

10.52 (a) $\text{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix}\right)$, $\quad \text{diag}\left(\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, [2]. \ [2], \begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix}\right)$;

(b) $\text{diag}\left(\begin{bmatrix} 7 & 1 \\ & 7 \end{bmatrix}, \begin{bmatrix} 7 & 1 \\ & 7 \end{bmatrix}, [7]\right)$, $\quad \text{diag}\left(\begin{bmatrix} 7 & 1 \\ & 7 \end{bmatrix}, [7], [7], [7]\right)$;

(c) Seja M_k um bloco de Jordan com $\lambda = 2$ de ordem k. Então $\text{diag}(M_3, M_3, M_1)$, $\text{diag}(M_3, M_2, M_2)$, $\text{diag}(M_3, M_2, M_1, M_1)$, $\text{diag}(M_3, M_1, M_1, M_1, M_1)$

10.60 Sejam $A = \begin{bmatrix} 0 & -3 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 8 \\ 1 & 0 & -12 \\ 0 & 1 & 6 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & -4 \\ 1 & 4 \end{bmatrix}$.

(a) $\text{diag}(A, A, B), \text{diag}(A, B, B), \text{diag}(A, B, -1, -1)$; (b) $\text{diag}(C, C), \text{diag}(C, D, 2), \text{diag}(C, 2, 2, 2)$

10.61 Sejam $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 3 \\ 1 & 0 \end{bmatrix}$.

(a) $\text{diag}(A, B)$, (b) $\text{diag}(A, \sqrt{3}, -\sqrt{3})$, (c) $\text{diag}(i, -i, \sqrt{3}, -\sqrt{3})$

10.62 A matriz companheira com última coluna dada por $[-\lambda^4, 4\lambda^3, -6\lambda^2, 4\lambda]^T$

Funcionais Lineares e o Espaço Dual

11.1 INTRODUÇÃO

Neste capítulo estudamos as transformações lineares de um espaço vetorial *V* em seu corpo de escalares *K*. (Salvo menção explícita em contrário, interpretamos *K* como um espaço vetorial sobre si mesmo.) Naturalmente, todos os teoremas e resultados obtidos para transformações arbitrárias de *V* continuam valendo nesse caso especial. Contudo, tratamos dessas transformações lineares em separado, em virtude de sua importância fundamental e porque a relação especial de *V* com *K* dá origem a novos conceitos e resultados que não se aplicam ao caso geral.

11.2 FUNCIONAIS LINEARES E O ESPAÇO DUAL

Seja *V* um espaço vetorial sobre o corpo *K*. Uma aplicação $\phi: V \to K$ é denominada *funcional linear* de *V* se, para quaisquer $u, v \in V$ e $a, b \in K$,

$$\phi(au + bv) = a\phi(u) + b\phi(v)$$

Em outras palavras, um funcional linear de *V* é uma transformação linear de *V* em *K*.

Exemplo 11.1

(a) Seja $\pi_i: K^n \to K$ a *i*-ésima aplicação projeção, isto é, $\pi_i(a_1, a_2, \ldots a_n) = a_i$. Então π_i é linear e, portanto, um funcional linear de K^n.

(b) Seja *V* o espaço vetorial dos polinômios em *t* sobre **R**. Seja $\mathbf{J}: V \to \mathbf{R}$ o operador integral definido por $\mathbf{J}(p(t)) = \int_0^1 p(t)\, dt$. Sabemos que **J** é linear e, portanto, um funcional linear de *V*.

(c) Seja *V* o espaço vetorial das matrizes quadradas de ordem *n* sobre *K*. Seja $T: V \to K$ a aplicação traço

$$T(A) = a_{11} + a_{22} + \cdots + a_{nn}, \qquad \text{com} \qquad A = [a_{ij}]$$

Ou seja, *T* associa a uma matriz *A* a soma de suas entradas diagonais. Essa aplicação é linear (Problema 11.24) e, portanto, um funcional linear de *V*.

Pelo Teorema 5.10, o conjunto dos funcionais lineares de um espaço vetorial *V* sobre *K* também é um espaço vetorial sobre *K* com adição e multiplicação por escalar definidas por

$$(\phi + \sigma)(v) = \phi(v) + \sigma(v) \qquad \text{e} \qquad (k\phi)(v) = k\phi(v)$$

onde ϕ e σ são funcionais lineares de *V* e $k \in K$. Esse espaço vetorial é denominado *espaço dual* de *V* e é denotado por *V**.

Exemplo 11.2 Seja $V = K^n$ o espaço vetorial das ênuplas de K, que escrevemos como vetores coluna. Então o espaço dual V^* pode ser identificado com o espaço dos vetores linha. Em particular, qualquer funcional linear $\phi = (a_1, \ldots, a_n)$ em V^* tem a representação

$$\phi(x_1, x_2, \ldots, x_n) = [a_1, a_2, \ldots, a_n][x_2, x_2, \ldots, x_n]^T = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

Historicamente, a expressão formal à direita dessa igualdade era denominada *forma linear*.

11.3 BASE DUAL

Suponha que V seja um espaço vetorial de dimensão n sobre K. Pelo Teorema 5.11, a dimensão do espaço dual V^* também é n (pois K tem dimensão 1 sobre si mesmo). Na verdade, cada base de V determina uma base de V^* como segue (ver Problema 11.3 para a demonstração).

Teorema 11.1 Seja $\{v_1, ..., v_n\}$ uma base de V sobre K. Sejam $\phi_1, \ldots, \phi_n \in V^*$ os funcionais lineares definidos por

$$\phi_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Então $\{\phi_1, \ldots, \phi_n\}$ é uma base de V^*.

A base $\{\phi_i\}$ desse teorema é denominada *base dual* de $\{v_i\}$. A fórmula dada na definição da base dual, utilizando o delta de Kronecker δ_{ij}, é uma maneira abreviada de escrever

$$\begin{array}{l} \phi_1(v_1) = 1, \;\; \phi_1(v_2) = 0, \;\; \phi_1(v_3) = 0, \;\; \ldots, \;\; \phi_1(v_n) = 0 \\ \phi_2(v_1) = 0, \;\; \phi_2(v_2) = 1, \;\; \phi_2(v_3) = 0, \;\; \ldots, \;\; \phi_2(v_n) = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \phi_n(v_1) = 0, \;\; \phi_n(v_2) = 0, \;\; \ldots, \phi_n(v_{n-1}) = 0, \;\; \phi_n(v_n) = 1 \end{array}$$

Pelo Teorema 5.2, essas transformações lineares ϕ_i são únicas e estão bem definidas.

Exemplo 11.3 Considere a base $\{v_1 = (2, 1), v_2 = (3, 1)\}$ de $\mathbf{R}^2$. Encontre a base dual $\{\phi_1, \phi_2\}$.

Queremos encontrar os funcionais lineares $\phi_1(x, y) = ax + by$ e $\phi_2(x, y) = cx + dy$ tais que

$$\phi_1(v_1) = 1, \qquad \phi_1(v_2) = 0, \qquad \phi_2(v_2) = 0, \qquad \phi_2(v_2) = 1$$

Essas quatro condições levam aos dois sistemas de equações lineares a seguir.

$$\left.\begin{array}{l} \phi_1(v_1) = \phi_1(2, 1) = 2a + b = 1 \\ \phi_1(v_2) = \phi_1(3, 1) = 3a + b = 0 \end{array}\right\} \quad \text{e} \quad \left.\begin{array}{l} \phi_2(v_1) = \phi_2(2, 1) = 2c + d = 0 \\ \phi_2(v_2) = \phi_2(3, 1) = 3c + d = 1 \end{array}\right\}$$

As soluções são $a = -1$, $b = 3$ e $c = 1$, $d = -2$. Logo, $\phi_1(x, y) = -x + 3y$ e $\phi_2(x, y) = x - 2y$ constituem a base dual de $\{v_1, v_2\}$.

Os dois teoremas seguintes (demonstrados nos Problemas 11.4 e 11.5, respectivamente) estabelecem relações entre bases e bases duais.

Teorema 11.2 Sejam $\{v_1, ..., v_n\}$ uma base de V e $\{\phi_1, \ldots, \phi_n\}$ a base dual de V^*. Valem as afirmações seguintes.

(i) Dado qualquer vetor $u \in V$, $u = \phi_1(u)v_1 + \phi_2(u)v_2 + \cdots + \phi_n(u)v_n$.
(ii) Dado qualquer funcional linear $\sigma \in V^*$, $\sigma = \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n$.

Teorema 11.3 Sejam $\{v_1, ..., v_n\}$ e $\{w_1, ..., w_n\}$ bases de V e sejam $\{\phi_1, \ldots, \phi_n\}$ e $\{\sigma_1, \ldots, \sigma_n\}$ as bases de V^* duais de $\{v_i\}$ e $\{w_i\}$, respectivamente. Seja P a matriz de mudança de base de $\{v_i\}$ para $\{w_i\}$. Então $(P^{-1})^T$ é a matriz de mudança de base de $\{\phi_i\}$ para $\{\sigma_i\}$.

11.4 ESPAÇO BIDUAL

Repetimos: cada espaço vetorial V tem um espaço dual V^*, que consiste em todos os funcionais lineares de V. Assim, V^* tem um espaço dual V^{**}, denominado *espaço bidual* de V, que consiste em todos os funcionais lineares de V^*.

Mostremos que cada $v \in V$ determina um elemento específico $\hat{v} \in V^{**}$. Dado qualquer $\phi \in V^*$, definimos

$$\hat{v}(\phi) = \phi(v)$$

Resta mostrar que essa aplicação $\hat{v}: V^* \to K$ é linear. Dados quaisquer escalares $a, b \in K$ e funcionais lineares $\phi, \sigma \in V^*$, temos

$$\hat{v}(a\phi + b\sigma) = (a\phi + b\sigma)(v) = a\phi(v) + b\sigma(v) = a\hat{v}(\phi) + b\hat{v}(\sigma)$$

Assim, $\hat{v}$ é linear e, portanto, $\hat{v} \in V^{**}$. Vale o teorema seguinte (demonstrado no Problema 12.7).

Teorema 11.4 Se V tem dimensão finita, então a aplicação $v \mapsto \hat{v}$ é um isomorfismo de V sobre V^{**}.

Essa aplicação $v \mapsto \hat{v}$ é denominada *aplicação natural* de V em V^{**}. Enfatizamos que a aplicação natural nunca é sobrejetora se V não for de dimensão finita. No entanto, ela sempre é linear e, mais que isso, injetora.

Suponha, agora, que V tenha dimensão finita. Pelo Teorema 11.4, a aplicação natural determina um isomorfismo entre V e V^{**}. Salvo menção explícita em contrário, identificamos V com V^{**} por meio desse isomorfismo. Dessa forma, consideramos V como o espaço dos funcionais lineares de V^* e escreveremos $V = V^{**}$. Observamos que se $\{\phi_i\}$ for a base de V^* que é dual da base $\{v_i\}$ de V, então $\{v_i\}$ é a base de $V^{**} = V$ que é dual da base $\{\phi_i\}$.

11.5 ANULADORES

Seja W um subconjunto (não necessariamente um subespaço) do espaço vetorial V. Dizemos que um funcional linear $\phi \in V^*$ é um *anulador* de W se $\phi(w) = 0$, para cada $w \in W$, ou seja, se $\phi(W) = \{0\}$. O conjunto de todos esses funcionais lineares é denotado por W^0 e denominado *anulador* de W. Mostremos que o anulador de W é um subespaço de V^*. Claramente, $0 \in W^0$. Sejam, agora, $\phi, \sigma \in W^0$. Então, dados quaisquer escalares $a, b \in K$ e qualquer $w \in W$, temos

$$(a\phi + b\sigma)(w) = a\phi(w) + b\sigma(w) = a0 + b0 = 0$$

e, portanto, $a\phi + b\sigma \in W^0$. Assim, W^0 é um subespaço de V^*.

No caso em que W é um subespaço de V, temos a relação seguinte entre W e seu anulador W^0 (ver Problema 11.11 para a demonstração).

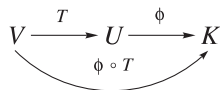
Teorema 11.5 Sejam V um espaço de dimensão finita e W um subespaço de V. Então

$$\text{(i) } \dim W + \dim W^0 = \dim V \qquad \text{e} \qquad \text{(ii) } W^{00} = W$$

Aqui, $W^{00} = \{v \in V : \phi(v) = 0 \text{ para cada } \phi \in W^0\}$ ou, equivalentemente, $W^{00} = (W^0)^0$, onde interpretamos W^{00} como um subespaço de V pela identificação de V com V^{**}.

11.6 TRANSPOSTA DE UMA TRANSFORMAÇÃO LINEAR

Seja $T: V \to U$ uma transformação linear qualquer de um espaço vetorial V num espaço vetorial U. Então, dado qualquer funcional linear $\phi \in U^*$, a composta $\phi \circ T$ é uma transformação linear de V em K.

$$V \xrightarrow{T} U \xrightarrow{\phi} K \qquad (V \xrightarrow{\phi \circ T} K)$$

Ou seja, $\phi \circ T \in V^*$. Assim, a correspondência

$$\phi \mapsto \phi \circ T$$

é uma aplicação de U^* em V^*, que denotamos por T^t e denominamos *transposta* de T. Em outras palavras, $T^t: U^* \to V^*$ é definida por

$T^t(\phi) = \phi \circ T$

Assim, $(T^t(\phi))(v) = \phi(T(v))$, para cada $v \in V$.

Teorema 11.6 A aplicação transposta T^t de T é uma transformação linear.

Demonstração: Dados quaisquer escalares $a, b \in K$ e funcionais lineares $\phi, \sigma \in U^*$,

$$T^t(a\phi + b\sigma) = (a\phi + b\sigma) \circ T = a(\phi \circ T) + b(\sigma \circ T) = aT^t(\phi) + bT^t(\sigma)$$

Assim, T^t é linear, como queríamos demonstrar.

Enfatizamos que, se T é uma transformação linear de V em U, então T^t é uma transformação linear de U^* em V^*. A nomenclatura "transposta" para a transformação T^t sem dúvida deriva da afirmação do teorema seguinte (demonstrado no Problema 11.16).

Teorema 11.7 Sejam $T: V \to U$ uma transformação linear e A a representação matricial de T em relação às bases $\{v_i\}$ de V e $\{u_i\}$ de U. Então a matriz transposta A^T é a representação matricial da transformação transposta $T^t: U^* \to V^*$ em relação às bases duais de $\{u_i\}$ e $\{v_i\}$.

Problemas Resolvidos

Espaços duais e bases duais

11.1 Encontre a base $\{\phi_1, \phi_2, \phi_3\}$ que é dual à base de $\mathbf{R}^3$ dada.

$$\{v_1 = (1, -1, 3), \quad v_2 = (0, 1, -1), \quad v_3 = (0, 3, -2)\}$$

Os funcionais lineares podem ser escritos da forma

$$\phi_1(x, y, z) = a_1x + a_2y + a_3z, \qquad \phi_2(x, y, z) = b_1x + b_2y + b_3z, \qquad \phi_3(x, y, z) = c_1x + c_2y + c_3z$$

Por definição de base dual, $\phi_i(v_j) = 0$ para $i \neq j$, mas $\phi_i(v_j) = 1$ para $i = j$. Obtemos ϕ_1 exigindo $\phi_1(v_1) = 1, \ \phi_1(v_2) = 0, \ \phi_1(v_3) = 0$. Isso fornece

$$\phi_1(1, -1, 3) = a_1 - a_2 + 3a_3 = 1, \qquad \phi_1(0, 1, -1) = a_2 - a_3 = 0, \qquad \phi_1(0, 3, -2) = 3a_2 - 2a_3 = 0$$

Resolvendo o sistema de equações, obtemos $a_1 = 1$, $a_2 = 0$, $a_3 = 0$. Assim, $\phi_1(x, y, z) = x$.

Obtemos ϕ_2 exigindo $\phi_2(v_1) = 0, \ \phi_2(v_2) = 1, \ \phi_2(v_3) = 0$. Isso fornece

$$\phi_2(1, -1, 3) = b_1 - b_2 + 3b_3 = 0, \qquad \phi_2(0, 1, -1) = b_2 - b_3 = 1, \qquad \phi_2(0, 3, -2) = 3b_2 - 2b_3 = 0$$

Resolvendo o sistema de equações, obtemos $b_1 = 7$, $b_2 = -2$, $a_3 = -3$. Assim, $\phi_2(x, y, z) = 7x - 2y - 3z$.

Obtemos ϕ_3 exigindo $\phi_3(v_1) = 0, \ \phi_3(v_2) = 0, \ \phi_3(v_3) = 1$. Isso fornece

$$\phi_3(1, -1, 3) = c_1 - c_2 + 3c_3 = 0, \qquad \phi_3(0, 1, -1) = c_2 - c_3 = 0, \qquad \phi_3(0, 3, -2) = 3c_2 - 2c_3 = 1$$

Resolvendo o sistema de equações, obtemos $c_1 = -2$, $c_2 = 1$, $c_3 = 1$. Assim, $\phi_3(x, y, z) = -2x + y + z$.

11.2 Seja $V = \{a + bt : a, b \in \mathbf{R}\}$ o espaço vetorial dos polinômios reais de grau ≤ 1. Encontre a base $\{v_1, v_2\}$ de V que é dual da base $\{\phi_1, \phi_2\}$ de V^* dada por

$$\phi_1(f(t)) = \int_0^1 f(t)\, dt \qquad \text{e} \qquad \phi_2(f(t)) = \int_0^2 f(t)\, dt$$

Sejam $v_1 = a + bt$ e $v_2 = c + dt$. Por definição de base dual,

$$\phi_1(v_1) = 1, \qquad \phi_1(v_2) = 0 \qquad \text{e} \qquad \phi_2(v_1) = 0, \qquad \phi_i(v_j) = 1$$

Assim,

$$\left.\begin{array}{l}\phi_1(v_1) = \int_0^1 (a+bt)\,dt = a + \frac{1}{2}b = 1 \\ \phi_2(v_1) = \int_0^2 (a+bt)\,dt = 2a + 2b = 0\end{array}\right\} \quad \text{e} \quad \left.\begin{array}{l}\phi_1(v_2) = \int_0^1 (c+dt)\,dt = c + \frac{1}{2}d = 0 \\ \phi_2(v_2) = \int_0^2 (c+dt)\,dt = 2c + 2d = 1\end{array}\right\}$$

Resolvendo cada sistema, obtemos $a = 2$, $b = -2$ e $c = -\frac{1}{2}$, $d = 1$. Assim, $\{v_1 = 2 - 2t,\ v_2 = -\frac{1}{2} + t\}$ é a base de V que é dual de $\{\phi_1, \phi_2\}$.

11.3 Demonstre o Teorema 11.1. Seja $\{v_1, ..., v_n\}$ uma base de V sobre K. Sejam $\phi_1, \ldots, \phi_n \in V^*$ os funcionais lineares definidos por $\phi_i(v_j) = 0$ com $i \neq j$, mas $\phi_i(v_j) = 1$ com $i = j$. Então $\{\phi_1, \ldots, \phi_n\}$ é uma base de V^*.

Começamos mostrando que $\{\phi_1, \ldots, \phi_n\}$ gera V^*. Seja ϕ um elemento arbitrário de V^* e suponha que

$$\phi(v_1) = k_1, \quad \phi(v_2) = k_2, \quad \ldots, \quad \phi(v_n) = k_n$$

Denote $\sigma = k_1\phi_1 + \cdots + k_n\phi_n$. Então

$$\begin{aligned}\sigma(v_1) &= (k_1\phi_1 + \cdots + k_n\phi_n)(v_1) = k_1\phi_1(v_1) + k_2\phi_2(v_1) + \cdots + k_n\phi_n(v_1) \\ &= k_1 \cdot 1 + k_2 \cdot 0 + \cdots + k_n \cdot 0 = k_1\end{aligned}$$

Analogamente, para $i = 2, ..., n$,

$$\sigma(v_i) = (k_1\phi_1 + \cdots + k_n\phi_n)(v_i) = k_1\phi_1(v_i) + \cdots + k_i\phi_i(v_i) + \cdots + k_n\phi_n(v_i) = k_i$$

Assim, $\phi(v_i) = \sigma(v_i)$, para $i = 1, ..., n$. Como ϕ e σ coincidem nos vetores da base, resulta $\phi = \sigma = k_1\phi_1 + \cdots + k_n\phi_n$. Por isso, $\{\phi_1, \ldots, \phi_n\}$ gera V^*.

Resta mostrar que $\{\phi_1, \ldots, \phi_n\}$ é linearmente independente. Suponha que

$$a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n = 0$$

Aplicando ambos lados no vetor v_1, obtemos

$$\begin{aligned}0 = 0(v_1) &= (a_1\phi_1 + \cdots + a_n\phi_n)(v_1) = a_1\phi_1(v_1) + a_2\phi_2(v_1) + \cdots + a_n\phi_n(v_1) \\ &= a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_n \cdot 0 = a_1\end{aligned}$$

Analogamente, para $i = 2, ..., n$,

$$0 = 0(v_i) = (a_1\phi_1 + \cdots + a_n\phi_n)(v_i) = a_1\phi_1(v_i) + \cdots + a_i\phi_i(v_i) + \cdots + a_n\phi_n(v_i) = a_i$$

Logo, $a_1 = 0, \ldots, a_n = 0$. Assim, $\{\phi_1, \ldots, \phi_n\}$ é linearmente independente e, portanto, uma base de V^*.

11.4 Demonstre o Teorema 11.2. Sejam $\{v_1, ..., v_n\}$ uma base de V e $\{\phi_1, \ldots, \phi_n\}$ a base dual de V^*. Dados quaisquer $u \in V$ e $\sigma \in V^*$, (i) $u = \sum_i \phi_i(u)v_i$, (ii) $\sigma = \sum_i \phi(v_i)\phi_i$.

Suponha que

$$u = a_1v_1 + a_2v_2 + \cdots + a_nv_n \tag{1}$$

Então

$$\phi_1(u) = a_1\phi_1(v_1) + a_2\phi_1(v_2) + \cdots + a_n\phi_1(v_n) = a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_n \cdot 0 = a_1$$

Analogamente, para $i = 2, ..., n$,

$$\phi_i(u) = a_1\phi_i(v_1) + \cdots + a_i\phi_i(v_i) + \cdots + a_n\phi_i(v_n) = a_i$$

Logo, $\phi_1(u) = a_1,\ \phi_2(u) = a_2, \ldots, \phi_n(u) = a_n$. Substituindo esses valores em (1), obtemos (i).

Para provar (ii), aplicamos o funcional σ a ambos lados de (1), obtendo

$$\begin{aligned}\sigma(u) &= \phi_1(u)\sigma(v_1) + \phi_2(u)\sigma(v_2) + \cdots + \phi_n(u)\sigma(v_n) \\ &= \sigma(v_1)\phi_1(u) + \sigma(v_2)\phi_2(u) + \cdots + \sigma(v_n)\phi_n(u) \\ &= (\sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n)(u)\end{aligned}$$

Como isso vale para cada $u \in V$, temos $\sigma = \sigma(v_1)\phi_2 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n$, como queríamos mostrar.

11.5 Demonstre o Teorema 11.3. Sejam $\{v_i\}$ e $\{w_i\}$ bases de V e sejam $\{\phi_i\}$ e $\{\sigma_i\}$ as respectivas bases duas de V^*. Seja P a matriz de mudança de base de $\{v_i\}$ para $\{w_i\}$. Então $(P^{-1})^T$ é a matriz de mudança de base de $\{\phi_i\}$ para $\{\sigma_i\}$.

Suponha que, para $i = 1, ..., n$,

$$w_i = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n \qquad \text{e} \qquad \sigma_i = b_{i1}\phi_1 + b_{i2}\phi_2 + \cdots + a_{in}v_n$$

Então $P = [a_{ij}]$ e $Q = [b_{ij}]$. Queremos provar que $Q = (P^{-1})^T$.

Denotemos a i-ésima linha de Q por R_i e a j-ésima coluna de P^T por C_j. Então

$$R_i = (b_{i1}, b_{i2}, \ldots, b_{in}) \qquad \text{e} \qquad C_j = (a_{j1}, a_{j2}, \ldots, a_{jn})^T$$

Por definição de base dual,

$$\sigma_i(w_j) = (b_{i1}\phi_1 + b_{i2}\phi_2 + \cdots + b_{in}\phi_n)(a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{jn}v_n)$$
$$= b_{i1}a_{j1} + b_{i2}a_{j2} + \cdots + b_{in}a_{jn} = R_iC_j = \delta_{ij}$$

onde δ_{ij} é o delta de Kronecker. Logo,

$$QP^T = [R_iC_j] = [\delta_{ij}] = I$$

Assim, $Q = (P^T)^{-1} = (P^{-1})^T$, como queríamos mostrar.

11.6 Suponha que $v \in V$, com $v \neq 0$ e dim $V = n$. Mostre que existe $\phi \in V^*$ tal que $\phi(v) \neq 0$.

Estendemos $\{v\}$ a uma base $\{v, v_2, ..., v_n\}$ de V. Pelo Teorema 5.2, existe uma única transformação linear $\phi: V \rightarrow K$ tal que $\phi(v) = 1$ e $\phi(v_i) = 0,\ i = 2, \ldots, n$. Assim, ϕ tem as propriedades procuradas.

11.7 Demonstre o Teorema 11.4. Se V tem dimensão finita, então a aplicação $v \mapsto \hat{v}$ é um isomorfismo de V sobre V^{**}.

Começamos provando que a aplicação $v \mapsto \hat{v}$ é linear, ou seja, que dados vetores $v, w \in V$ e escalares $a, b \in K$, quaisquer, $\widehat{av + bw} = a\hat{v} + b\hat{w}$. Dado qualquer funcional linear $\phi \in V^*$,

$$\widehat{av + bw}(\phi) = \phi(av + bw) = a\phi(v) + b\phi(w) = a\hat{v}(\phi) + b\hat{w}(\phi) = (a\hat{v} + b\hat{w})(\phi)$$

Como $\widehat{av + bw}(\phi) = (a\hat{v} + b\hat{w})(\phi)$ para cada $\phi \in V^*$, temos $\widehat{av + bw} = a\hat{v} + b\hat{w}$. Assim, $v \mapsto \hat{v}$ é uma transformação linear.

Suponha, agora, que $v \in V$, $v \neq 0$. Então, pelo Problema 11.6, existe $\phi \in V^*$ para o qual $\phi(v) \neq 0$. Logo, $\hat{v}(\phi) = \phi(v) \neq 0$ e, assim, $\hat{v} \neq 0$. Como $v \neq 0$ implica $\hat{v} \neq 0$, a transformação linear $v \mapsto \hat{v}$ é não singular.

Ocorre que dim V = dim V^* = dim V^{**}, já que V tem dimensão finita. Pelo Teorema 5.9, a transformação $v \mapsto \hat{v}$ é um isomorfismo de V sobre V^{**}.

Anuladores

11.8 Mostre que se $\phi \in V^*$ anula um subconjunto S de V, então ϕ anula o espaço gerado ger(S) por S. Assim, $S^0 = [\text{ger}(S)]^0$.

Seja $v \in$ ger(S). Então existem $w_1, \ldots, w_r \in S$ para os quais $v = a_1w_1 + a_2w_2 + \cdots + a_rw_r$. Temos

$$\phi(v) = a_1\phi(w_1) + a_2\phi(w_2) + \cdots + a_r\phi(w_r) = a_1 0 + a_2 0 + \cdots + a_r 0 = 0$$

Como v é um elemento arbitrário de ger(S), ϕ anula ger(S), como queríamos mostrar.

11.9 Encontre uma base do anulador W^0 do subespaço W de $\mathbf{R}^4$ gerado por

$$v_1 = (1, 2, -3, 4) \qquad \text{e} \qquad v_2 = (0, 1, 4, -1)$$

Pelo Problema 11.8, basta encontrar uma base do conjunto de funcionais lineares ϕ tais que $\phi(v_1) = 0$ e $\phi(v_2) = 0$, sendo $\phi(x_1, x_2, x_3, x_4) = ax_1 + bx_2 + cx_3 + dx_4$. Assim,

$$\phi(1, 2, -3, 4) = a + 2b - 3c + 4d = 0 \qquad \text{e} \qquad \phi(0, 1, 4, -1) = b + 4c - d = 0$$

O sistema de duas equações nas incógnitas a, b, c, d está em forma escalonada, com variáveis livres c e d.

(1) Tomando $c = 1, d = 0$, obtemos a solução $a = 11, b = -4, c = 1, d = 0$.

(2) Tomando $c = 0, d = 1$, obtemos a solução $a = 6, b = -1, c = 0, d = 0$.

Os funcionais lineares $\phi_1(x_i) = 11x_1 - 4x_2 + x_3$ e $\phi_2(x_i) = 6x_1 - x_2 + x_4$ formam uma base de W^0.

11.10 Mostre as afirmações. (a) Dado qualquer subconjunto S de V, $S \subseteq S^{00}$. (b) Se $S_1 \subseteq S_2$, então $S_2^0 \subseteq S_1^0$.

(a) Seja $v \in S$. Então, dado qualquer funcional linear $\phi \in S^0$, $\hat{v}(\phi) = \phi(v) = 0$. Logo, $\hat{v} \in (S^0)^0$. Assim, pela identificação de V com V**, temos $v \in S^{00}$. Por isso, $S \subseteq S^{00}$.

(b) Seja $\phi \in S_2^0$. Então $\phi(v) = 0$, para cada $v \in S_2$. Mas $S_1 \subseteq S_2$, portanto, ϕ anula cada elemento de S_1 (ou seja, $\phi \in S_1^0$). Assim, $S_2^0 \subseteq S_1^0$.

11.11 Demonstre o Teorema 11.5. Sejam V um espaço de dimensão finita e W um subespaço de V. Então
(i) $\dim W + \dim W^0 = \dim V$, (ii) $W^{00} = W$.

(i) Suponha que $\dim V = n$ e que $W = r \leq n$. Queremos mostrar que $\dim W^0 = n - r$. Escolhemos uma base $\{w_1, ..., w_r\}$ de W e estendemos essa base a uma base de V, digamos, $\{w_1, \dots, w_r, v_1, \dots, v_{n-r}\}$. Considere a base dual

$$\{\phi_1, \dots, \phi_r, \sigma_1, \dots, \sigma_{n-r}\}$$

Por definição de base dual, cada um dos σ's dessa base anula cada w_i, de modo que $\sigma_1, \dots, \sigma_{n-r} \in W^0$. Afirmamos que $\{\sigma_i\}$ é uma base de W^0. Agora, $\{\sigma_j\}$ é parte de uma base de V^*, portanto é linearmente independente.

Em seguida mostramos que $\{\phi_j\}$ gera W^0. Seja $\sigma \in W^0$. Pelo Teorema 11.2,

$$\begin{aligned}\sigma &= \sigma(w_1)\phi_1 + \cdots + \sigma(w_r)\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} \\ &= 0\phi_1 + \cdots + 0\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} \\ &= \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r}\end{aligned}$$

Consequentemente, $\{\sigma_1, \dots, \sigma_{n-r}\}$ gera W^0 e é, portanto, uma base de W^0. Em vista disso, como queríamos,

$$\dim W^0 = n - r = \dim V - \dim W.$$

(ii) Suponha que $\dim V = n$ e que $\dim W = r$. Então $\dim V^* = n$ e, por (i), $\dim W^0 = n - r$. Assim, por (i) temos $\dim W^{00} = n - (n - r) = r$ e, portanto, $\dim W = \dim W^{00}$. Pelo Problema 11.10, $W \subseteq W^{00}$. Em vista disso, $W = W^{00}$.

11.12 Sejam U e W subespaços de V. Prove que $(U + W)^0 = U^0 \cap W^0$.

Seja $\phi \in (U + W)^0$. Então ϕ anula $U + W$ e, portanto, em particular, ϕ anula U e, também, W. Logo, $\phi \in U^0$ e $\phi \in W^0$, de modo que $\phi \in U^0 \cap W^0$. Assim, $(U + W)^0 \subseteq U^0 \cap W^0$.
Por outro lado, suponha que $\sigma \in U^0 \cap W^0$. Então σ anula U e, também, W. Se $v \in U + W$, então $v = u + w$, com $u \in U$ e $w \in W$. Logo, $\sigma(v) = \sigma(u) + \sigma(w) = 0 + 0 = 0$. Assim, σ anula $U + W$, ou seja, $\sigma \in (U + W)^0$. Em vista disso, $U^0 + W^0 \subseteq (U + W)^0$.
Juntando as duas inclusões, obtemos o resultado procurado.

OBSERVAÇÃO Observe que não utilizamos argumento dimensional nessa prova, o que significa que o resultado vale em espaços de dimensão finita ou infinita.

Transposta de uma transformação linear

11.13 Seja ϕ o funcional linear de $\mathbf{R}^2$ definido por $\phi(x, y) = x - 2y$. Para cada um dos operadores lineares T de $\mathbf{R}^2$ dados, calcule $(T^t(\phi))(x, y)$.
(a) $T(x, y) = (x, 0)$, (b) $T(x, y) = (y, \ x + y)$, (c) $T(x, y) = (2x - 3y, \ 5x + 2y)$

Por definição, $T^t(\phi) = \phi \circ T$, isto é, $(T^t(\phi))(v) = \phi(T(v))$ para cada v. Logo,

(a) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(x, 0) = x$

(b) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(y, \ x + y) = y - 2(x + y) = -2x - y$

(c) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(2x - 3y, \ 5x + 2y) = (2x - 3y) - 2(5x + 2y) = -8x - 7y$

11.14 Sejam $T: V \to U$ uma transformação linear e $T^t: U^* \to V^*$ sua transposta. Mostre que o núcleo de T^t é o anulador da imagem de T, isto é, $\text{Nuc}\, T^t = (\text{Im}\, T)^0$.

Seja $\phi \in \text{Nuc}\, T^t$, isto é, $T^t(\phi) = \phi \circ T = 0$. Se $u \in \text{Im}\, T$, então $u = T(v)$, para algum $v \in V$. Logo,

$$\phi(u) = \phi(T(v)) = (\phi \circ T)(v) = 0(v) = 0$$

Temos $\phi(u) = 0$, para cada $u \in \text{Im } T$. Logo, $\phi \in (\text{Im } T)^0$ e, assim, Nuc $T^t \subseteq (\text{Im } T)^0$.
Por outro lado, seja $\sigma \in (\text{Im } T)^0$, isto é, $\sigma(\text{Im } T) = \{0\}$. Então, para cada $v \in V$,

$$(T^t(\sigma))(v) = (\sigma \circ T)(v) = \sigma(T(v)) = 0 = 0(v)$$

Temos $(T^t(\sigma))(v) = 0(v)$, para cada $v \in V$. Logo, $T^t(\sigma) = 0$. Assim, $\sigma \in$ Ker T^t e, portanto, $(\text{Im } T)^0 \subseteq$ Ker T^t.
Juntando as duas inclusões, obtemos o resultado procurado.

11.15 Seja $T: V \to U$ uma transformação linear entre os espaços V e U de dimensão finita. Mostre que pos(T) = pos(T^t).

Digamos que dim $V = n$, dim $U = m$ e pos(T)= r. Pelo Teorma 11.5,

$$\dim(\text{Im } T)^0 = \dim u - \dim(\text{Im } T) = m - \text{ pos}(T) = m - r$$

Pelo Problema 11.4, $T^t = (\text{Im } T)^0$. Assim, $(T^t) = m - r$ e, portnto, como queríamos mostrar,

$$\text{pos}(T^t) = \dim U^* - \text{nul}(T^t) = m - (m - r) = r = \text{ pos}(T)$$

11.16 Demonstre o Teorema 11.7. Sejam $T: V \to U$ uma transformação linear e A a representação matricial de T em relação às bases $\{v_i\}$ de V e $\{u_i\}$ de U. Então a matriz transposta A^T é a representação matricial da transformação transposta $T^t: U^* \to V^*$ em relação às bases duais de $\{u_i\}$e$\{v_i\}$.

Digamos que, para $j = 1, ..., m$,

$$T(v_j) = a_{j1}u_1 + a_{j2}u_2 + \cdots + a_{jn}u_n \tag{1}$$

Queremos provar que, para $i = 1, ..., n$,

$$T^t(\sigma_i) = a_{1i}\phi_1 + a_{2i}\phi_2 + \cdots + a_{mi}\phi_m \tag{2}$$

em que $\{\sigma_i\}$ e $\{\phi_j\}$ são as bases duais de $\{u_i\}$ e $\{v_j\}$,, respectivamente.
Seja $v \in V$ dado e suponha que $v = k_1 v_1 + k_2 v_2 + \cdots + k_m v_m$. Então, por (1),

$$\begin{aligned}
T(v) &= k_1 T(v_1) + k_2 T(v_2) + \cdots + k_m T(v_m) \\
&= k_1(a_{11}u_1 + \cdots + a_{1n}u_n) + k_2(a_{21}u_1 + \cdots + a_{2n}u_n) + \cdots + k_m(a_{m1}u_1 + \cdots + a_{mn}u_n) \\
&= (k_1 a_{11} + k_2 a_{21} + \cdots + k_m a_{m1})u_1 + \cdots + (k_1 a_{1n} + k_2 a_{2n} + \cdots + k_m a_{mn})u_n \\
&= \sum_{i=1}^{n}(k_1 a_{1i} + k_2 a_{2i} + \cdots + k_m a_{mi})u_i
\end{aligned}$$

Logo, para $j = 1, ..., n$,

$$\begin{aligned}
(T^t(\sigma_j)(v)) = \sigma_j(T(v)) &= \sigma_j\left(\sum_{i=1}^{n}(k_1 a_{1i} + k_2 a_{2i} + \cdots + k_m a_{mi})u_i\right) \\
&= k_1 a_{1j} + k_2 a_{2j} + \cdots + k_m a_{mj}
\end{aligned} \tag{3}$$

Por ouro lado, para $j = 1, ..., n$,

$$\begin{aligned}
(a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(v) &= (a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(k_1 v_1 + k_2 v_2 + \cdots + k_m v_m) \\
&= k_1 a_{1j} + k_2 a_{2j} + \cdots + k_m a_{mj}
\end{aligned} \tag{4}$$

Como $v \in V$ é um vetor arbitrário, (3) e (4) implicam que

$$T^t(\sigma_j) = a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m, \qquad j = 1, \ldots, n$$

que é (2). Assim, o teorema está provado.

Problemas Complementares

Espaços duais e bases duais

11.17 Encontre (a) $\phi + \sigma$, (b) 3ϕ, (c) $2\phi - 5\sigma$, com $\phi:\mathbf{R}^3 \to \mathbf{R}$ e $\sigma:\mathbf{R}^3 \to \mathbf{R}$ definidos por

$$\phi(x, y, z) = 2x - 3y + z \qquad \text{e} \qquad \sigma(x, y, z) = 4x - 2y + 3z$$

11.18 Encontre a base dual de cada uma das bases de $\mathbf{R}^3$ dadas.
(a) $\{(1, 0, 0), \ (0, 1, 0), \ (0, 0, 1)\}$ (b) $\{(1, -2, 3), \ (1, -1, 1), \ (2, -4, 7)\}$.

11.19 Seja V o espaço vetorial dos polinômios sobre $\mathbf{R}$ de grau ≤ 2. Sejam ϕ_1, ϕ_2, ϕ_3 os funcionais lineares definidos por

$$\phi_1(f(t)) = \int_0^1 f(t)\, dt, \qquad \phi_2(f(t)) = f'(1), \qquad \phi_3(f(t)) = f(0)$$

Aqui, $f(t) = a + bt + ct^2 \in V$ e $f'(t)$ denota a derivada de $f(t)$. Encontre a base $\{f_1(t), f_2(t), f_3(t)\}$ de V que é dual de $\{\phi_1, \phi_2, \phi_3\}$.

11.20 Sejam $u, v \in V$ tais que $\phi(u) = 0$ implica $\phi(v) = 0$, para cada $\phi \in V^*$. Mostre que $v = ru$, para algum escalar r.

11.21 Sejam $\phi, \sigma \in V^*$ tais que $\phi(v) = 0$ implica $\sigma(v) = 0$, para cada $v \in V$. Mostre que $\sigma = k\phi$, para algum escalar k.

11.22 Seja V o espaço vetorial dos polinômios sobre K. Dado $a \in K$, defina $\phi_a:V \to K$ por $\phi_a(f(t)) = f(a)$. Mostre que (a) ϕ_a é linear, (b) se $a \neq b$, então $\phi_a \neq \phi_b$.

11.23 Seja V o espaço vetorial dos polinômios de grau ≤ 2. Sejam $a, b, c \in K$ escalares distintos. Sejam ϕ_a, ϕ_b, ϕ_c os funcionais lineares definidos por $\phi_a(f(t)) = f(a)$, $\phi_b(f(t)) = f(b)$, $\phi_c(f(t)) = f(c)$. Mostre que $\{\phi_a, \phi_b, \phi_c\}$ é linearmente independente e encontre a base $\{f_1(t), f_2(t), f_3(t)\}$ de V que é dual.

11.24 Seja V o espaço vetorial das matrizes quadradas de ordem n. Seja $T:V \to K$ a aplicação traço, ou seja, a aplicação $T(A) = a_{11} + a_{22} + \cdots + a_{nn}$, com $A = [a_{ij}]$. Mostre que T é linear.

11.25 Seja W um subespaço de V. Mostre que, para cada funcional linear ϕ de W existe um funcional linear σ de V tal que $\sigma(w) = \phi(w)$, para cada $w \in W$, ou seja, ϕ é a restrição de σ a W.

11.26 Seja $\{e_1, ..., e_n\}$ a base canônica de K^n. Mostre que a base dual é $\{\pi_1, \ldots, \pi_n\}$, em que π_i é a i-ésima projeção de K^n, ou seja, $\pi_i(a_1, \ldots, a_n) = a_i$.

11.27 Seja V um espaço vetorial sobre $\mathbf{R}$. Let $\phi_1, \phi_2 \in V^*$. Sejam $\sigma:V \to \mathbf{R}$ e suponha que $\sigma(v) = \phi_1(v)\phi_2(v)$, definida por V^*, também pertence a V^*. Mostre que ou $\phi_1 = 0$ ou $\phi_2 = 0$.

Anuladores

11.28 Seja W o subespaço de $\mathbf{R}^4$ gerado por $(1, 2, -3, 4)$, $(1, 3, -2, 6)$, $(1, 4, -1, 8)$. Encontre uma base do anulador de W.

11.29 Seja W o subespaço de $\mathbf{R}^3$ gerado por $(1, 1, 0)$ e $(0, 1, 1)$. Encontre uma base do anulador de W.

11.30 Mostre que, para cada subconjunto S de V, ger$(S) = S^{00}$, onde ger(S) denota o espaço gerado por S.

11.31 Sejam U e W subespaços de um espaço V de dimensão finita. Prove que $(U \cap W)^0 = U^0 + W^0$.

11.32 Suponha que $V = U \oplus W$. Prove que $V^0 = U^0 \oplus W^0$.

Transposta de uma transformação linear

11.33 Seja ϕ o funcional linear de $\mathbf{R}^2$ definido por $\phi(x, y) = 3x - 2y$. Para cada uma das transformações lineares $T:\mathbf{R}^3 \to \mathbf{R}^2$ dadas, calcule $(T^t(\phi))(x, y, z)$.
(a) $T(x, y, z) = (x + y, \ y + z)$, (b) $T(x, y, z) = (x + y + z, \ 2x - y)$

11.34 Suponha que $T_1: U \to V$ e $T_2: V \to W$ sejam lineares. Prove que $(T_2 \circ T_1)^t = T_1^t \circ T_2^t$.

11.35 Suponha que $T: V \to U$ seja linear e que V tenha dimensão finita. Prove que Im T^t = (Nuc T)^0.

11.36 Sejam $T: V \to U$ linear e $u \in U$ dados. Mostre que ou $u \in \text{Im } T$ ou existe $\phi \in V^*$ tal que $T^t(\phi) = 0$ e $\phi(u) = 1$.

11.37 Seja V de dimensão finita. Mostre que a aplicação $T \mapsto T^t$ é um isomorfismo de Hom(V, V) sobre Hom(V^*, V^*). (Aqui, T é qualquer operador linear de V.)

Problemas variados

11.38 Seja V um espaço vetorial sobre **R**. Definimos o *segmento de reta* $\overline{uv}$ ligando os pontos $u, v \in V$ por $\overline{uv} = \{tu + (1-t)v : 0 \leq t \leq 1\}$. Dizemos que um subconjunto S de V é *convexo* se $u, v \in S$ implica $\overline{uv} \subseteq S$. Dado $\phi \in V^*$, defina

$$W^+ = \{v \in V : \phi(v) > 0\}, \qquad W = \{v \in V : \phi(v) = 0\}, \qquad W^- = \{v \in V : \phi(v) < 0\}$$

Prove que W^+, W e W^- são convexos.

11.39 Seja V um espaço vetorial de dimensão finita. Um *hiperplano* H de V pode ser definido como o núcleo de um funcional linear não nulo ϕ de V. Mostre que cada subespaço de V é a interseção de um número finito de hiperplanos de V.

Respostas dos Problemas Complementares

11.17 (a) $6x - 5y + 4z$, (b) $6x - 9y + 3z$, (c) $-16x + 4y - 13z$

11.18 (a) $\phi_1 = x,\ \phi_2 = y,\ \phi_3 = z$; (b) $\phi_1 = -3x - 5y - 2z,\ \phi_2 = 2x + y,\ \phi_3 = x + 2y + z$

11.19 $f_1(t) = 3t - \frac{3}{2}t^2,\quad f_2(t) = -\frac{1}{2}t + \frac{3}{4}t^2,\quad f_3(t) = 1 - 3t + \frac{3}{2}t^2$

11.22 (b) Seja $f(t) = t$. Então $\phi_a(f(t)) = a \neq b = \phi_b(f(t))$ e, portanto, $\phi_a \neq \phi_b$.

11.23 $\left\{f_1(t) = \dfrac{t^2 - (b+c)t + bc}{(a-b)(a-c)},\ f_2(t) = \dfrac{t^2 - (a+c)t + ac}{(b-a)(b-c)},\ f_3(t) = \dfrac{t^2 - (a+b)t + ab}{(c-a)(c-b)}\right\}$

11.28 $\{\phi_1(x,y,z,t) = 5x - y + z,\ \ \phi_2(x,y,z,t) = 2y - t\}$

11.29 $\{\phi(x,y,z) = x - y + z\}$

11.33 (a) $(T^t(\phi))(x,y,z) = 3x + y - 2z$, (b) $(T^t(\phi))(x,y,z) = -x + 5y + 3z$

Formas Bilineares, Quadráticas e Hermitianas

12.1 INTRODUÇÃO

Neste capítulo generalizamos os conceitos de transformações lineares e funcionais lineares. Especificamente, introduzimos o conceito de uma forma bilinear. Essas formas bilineares também dão origem às formas quadráticas e hermitianas. Embora as formas quadráticas já tenham sido discutidas anteriormente, neste capítulo tratamos dessas formas independentemente dos resultados anteriores.

Embora o corpo K seja arbitrário, mais adiante nos restringiremos aos casos $K = \mathbf{R}$ e $K = \mathbf{C}$. Além disso, às vezes podemos precisar dividir por 2. Nesses casos, devemos supor que $1 + 1 \neq 0$, o que é válido em $K = \mathbf{R}$ ou $K = \mathbf{C}$.

12.2 FORMAS BILINEARES

Seja V um espaço vetorial de dimensão finita sobre um corpo K. Uma *forma bilinear* de V é uma aplicação $f : V \times V \to K$ tal que, dados $a, b \in K$ e $u_i, v_i \in V$ quaisquer,

(i) $f(au_1 + bu_2, \; v) = af(u_1, v) + bf(u_2, v)$,

(ii) $f(u, \; av_1 + bv_2) = af(u, v_1) + bf(u, v_2)$

A condição (i) significa que f é *linear na primeira variável* e a condição (ii) que f é *linear na segunda variável*.

Exemplo 12.1

(a) Seja f o produto escalar de $\mathbf{R}^n$, isto é, dados $u = (a_i)$ e $v = (b_i)$,

$$f(u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Então f é uma forma bilinear de $\mathbf{R}^n$. (Na verdade, qualquer produto interno sobre um espaço vetorial real V é uma forma bilinear de V.)

(b) Sejam ϕ e σ funcionais lineares arbitrários de V. Seja $f : V \times V \to K$ definida por $f(u, v) = \phi(u)\sigma(v)$. Então f é uma forma bilinear, já que ϕ e σ são, ambas, lineares.

(c) Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n qualquer sobre um corpo K. Então A pode ser identificada com a forma bilinear F de K^n dada a seguir, em que $X = [x_i]$ e $Y = [y_i]$ são vetores coluna de variáveis.

$$f(X, Y) = X^T A Y = \sum_{i,j} a_{ij} x_i y_i = a_{11} x_1 y_1 + a_{12} x_1 y_2 + \cdots + a_{nn} x_n y_n$$

Essa expressão formal nas variáveis x_i e y_i é denominada *polinômio bilinear* correspondente à matriz A. A equação (1) (ver adiante) mostra que, num certo sentido, toda forma bilinear é desse tipo.

Espaço das formas bilineares

Seja $B(V)$ o conjunto de todas as formas bilineares de V. $B(V)$ tem uma estrutura de espaço vetorial definindo, para quaisquer $f, g \in B(V)$ e $k \in K$, $f + g$ e kf por

$$(f+g)(u,v) = f(u,v) + g(u,v) \qquad \text{e} \qquad (kf)(u,v) = kf(u,v)$$

Vale o teorema seguinte (demonstrado no Problema 12.4).

Teorema 12.1 Seja V um espaço vetorial de dimensão n sobre K. Seja $\{\phi_1, \dots, \phi_n\}$ uma base qualquer do espaço dual V^*. Então $\{f_{ij} : i, j = 1, ..., n\}$ é uma base de $B(V)$, em que f_{ij} é definida por $f_{ij}(u,v) = \phi_i(u)\phi_j(v)$. Em particular, $\dim B(V) = n^2$.

12.3 FORMAS BILINEARES E MATRIZES

Sejam f uma forma bilinear de V e $S = \{u_1, ..., u_n\}$ uma base de V. Dados $u, v \in V$, temos

$$u = a_1u_1 + \cdots + a_nu_n \qquad \text{e} \qquad v = b_1u_1 + \cdots + b_nu_n$$

Então

$$f(u,v) = f(a_1u_1 + \cdots + a_nu_n,\ \ b_1u_1 + \cdots + b_nu_n) = \sum_{i,j} a_ib_jf(u_i,u_j)$$

Assim, f é completamente determinada pelos n^2 valores $f(u_i, u_j)$.

A matriz $A = [a_{ij}]$, em que $a_{ij} = f(u_i, u_j)$, é denominada *representação matricial* de f em relação à base S ou, simplesmente, a “matriz de f em S”. Essa matriz “representa” f no sentido seguinte; dados $u, v \in V$ quaisquer,

$$f(u,v) = \sum_{i,j} a_ib_j f(u_i,u_j) = [u]_S^T A [v]_S \tag{1}$$

[Como de hábito, $[u]_S$ denota o vetor (coluna) de coordenadas de u na base S.]

Mudança de base, matrizes congruentes

Agora perguntamos: como se transforma a matriz que representa uma forma bilinear quando escolhemos uma nova base? A resposta é dada no teorema seguinte (demonstrado no Problema 12.5).

Teorema 12.2 Seja P a matriz de mudança de base de uma base S para uma outra base S'. Se A é a representação matricial da forma bilinear f em relação à base original S, então $B = P^TAP$ é a representação matricial de f em relação à nova base S'.

Esse teorema motiva uma definição.

DEFINIÇÃO Dizemos que uma matriz B é *congruente* a uma matriz A (e escrevemos $B \simeq A$) se existir uma matriz não singular P tal que $B = P^TAP$.

Assim, o Teorema 12.2 afirma que são congruentes todas as matrizes que representam uma mesma forma bilinear. Observamos que matrizes congruentes têm o mesmo posto, pois P e P^T são não singulares; em particular, o conceito introduzido a seguir está bem definido.

DEFINIÇÃO O *posto* de uma forma bilinear f de V, denotado pos(f), é o posto de qualquer representação matricial de f. Dizemos que f é *degenerada* ou *não degenerada* dependendo de termos pos(f) < dim V ou pos(f) = dim V, respectivamente.

12.4 FORMAS BILINEARES ALTERNADAS

Seja f uma forma bilinear de V. Dizemos que f é

(i) *alternada* se $f(v,v) = 0$, para cada $v \in V$;

(ii) *antissimétrica* se $f(u,v) = -f(v,u)$, para quaisquer $u, v \in V$.

Agora, suponha que (i) seja verdadeira. Então (ii) também é verdadeira, pois, dados quaisquer $u, v \in V$, temos

$$0 = f(u+v, \quad u+v) = f(u,u) + f(u,v) + f(v,u) + f(v,v) = f(u,v) + f(v,u)$$

Por outro lado, suponha que (ii) seja verdadeira e, também, que $1+1 \neq 0$. Então (i) também é verdadeira, pois, dado qualquer $v \in V$, temos $f(u, v) = -f(v, u)$. Em outras palavras, ser alternada e ser antissimétrica são equivalentes, se $1+1 \neq 0$.

O teorema principal sobre a estrutura das formas bilineares alternadas (demonstrado no Problema 12.23) é dado a seguir.

Teorema 12.3 Seja f uma forma bilinear alternada de V. Então existe uma base de V na qual f é representada por uma matriz diagonal em blocos M do tipo

$$M = \operatorname{diag}\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \ldots, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad [0], \quad [0], \quad \ldots \quad [0] \right)$$

Além disso, o número de blocos não nulos é determinado de modo único por f [porque é igual à metade do posto de f].

Em particular, esse teorema mostra que, necessariamente, toda forma bilinear alternada tem posto par.

12.5 FORMAS BILINEARES SIMÉTRICAS, FORMAS QUADRÁTICAS

Nesta seção investigamos os conceitos importantes de formas bilineares simétricas e formas quadráticas e sua representação por meio de matrizes simétricas. A única restrição sobre o corpo K é que $1+1 \neq 0$. Na Seção 12.6 nos restringimos ao corpo real **R**, o que fornece resultados especiais importantes.

Formas bilineares simétricas

Seja f uma forma bilinear de V. Dizemos que f é *simétrica* se, dados quaisquer $u, v \in V$, temos

$$f(u, v) = f(v, u),$$

É fácil mostrar que f é simétrica se, e só se, qualquer representação matricial A de f é uma matriz simétrica.

O principal resultado para formas bilineares simétricas (demonstrado no Problema 12.10) é dado a seguir. (Enfatizamos que estamos supondo $1+1 \neq 0$.)

Teorema 12.4 Seja f uma forma bilinear simétrica de V. Então V tem uma base $\{v_1, ..., v_n\}$ na qual f é representada por uma matriz diagonal, isto é, na qual $f(v_i, v_j) = 0$, se $i \neq j$.

Teorema 12.4 (Forma Alternativa) Seja A uma matriz simétrica sobre K. Então A é congruente a uma matriz diagonal, ou seja, existe uma matriz não singular P tal que P^TAP é diagonal.

Algoritmo de diagonalização

Sabemos que uma matriz não singular P é o produto de matrizes elementares (ver Teorema 3.17). Dessa forma, uma maneira de obter a forma diagonal $D = P^TAP$ é por meio de uma sequência de operações elementares com as linhas e essa mesma sequência de operações elementares com as colunas. Aplicadas na matriz identidade I, essa mesma sequência de operações elementares fornece P^T. Podemos formalizar esse algoritmo como segue.

Algoritmo 12.1 (Diagonalização por Congruência de uma Matriz Simétrica) É dada uma matriz simétrica $A = [a_{ij}]$ de ordem n.

Passo 1 Forme a matriz $n \times 2n$ em blocos $M = [A_1, I]$, em que $A_1 = A$ é a metade à esquerda de M e a matriz identidade I de ordem n é a metade à direita de M.

Passo 2 Examine a entrada a_{11}. Há três casos.

Caso I $a_{11} \neq 0$ (Use a_{11} como um pivô para colocar zeros abaixo de a_{11} em M e à direita de a_{11} em A_1.) Para $i = 2, ..., n$,

(a) Aplique a operação com as linhas "Substituir R_i por $-a_{i1}R_1 + a_{11}R_i$".
(b) Aplique a operação correspondente com as colunas "Substituir C_i por $-a_{i1}C_1 + a_{11}C_i$".

Essas operações reduzem a matriz M à forma

$$M \sim \begin{bmatrix} a_{11} & 0 & * & * \\ 0 & A_2 & * & * \end{bmatrix} \qquad (*)$$

Caso II $a_{11} = 0$, mas $a_{kk} \neq 0$, para algum $k > 1$.

(a) Aplique a operação com as linhas "Trocar R_1 e R_k entre si".
(b) Aplique a operação correspondente com as colunas "Trocar C_1 e C_k entre si".

(Essas operações levam a_{kk} para a primeira entrada diagonal, que reduz a matriz ao Caso I.)

Caso III Todas entradas diagonais $a_{ii} = 0$, mas algum $a_{ij} \neq 0$.

(a) Aplique a operação com as linhas "Substituir R_i por $R_j + R_i$".
(b) Aplique a operação correspondente com as colunas "Substituir C_i por $C_j + C_i$".

(Essas operações levam $2a_{ij}$ para a i-ésima entrada diagonal, que reduz a matriz ao Caso II.)

Assim, finalmente, M é reduzida à forma (*), em que A_2 é uma matriz simétrica de ordem menor do que a ordem de A.

Passo 3 Repita o Passo 2 com cada nova matriz A_k (suprimindo as primeiras linha e coluna da matriz precedente) até que A esteja diagonalizada. Então, M está transformada à forma $M' = [D, Q]$, em que D é diagonal.

Passo 4 Tome $P = Q^T$. Então $D = P^TAP$.

OBSERVAÇÃO 1 Enfatizamos que, no Passo 2, as operações com as linhas alteram os dois lados de M, mas as operações com as colunas alteram apenas o lado esquerdo de M.

OBSERVAÇÃO 2 A condição $1 + 1 \neq 0$ é utilizada no Caso III, em que concluímos $2a_{ij} \neq 0$ a partir de $a_{ij} \neq 0$.

No Problema 12.9 apresentamos uma justificativa para esse algoritmo.

Exemplo 12.2 Seja $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{bmatrix}$. Aplique o Algoritmo 12.1 para encontrar uma matriz não singular P tal que $D = P^TAP$ seja diagonal.

Começamos construindo a matriz $M = [A, I]$, isto é, tomamos

$$M = [A, I] = \left[\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right]$$

Aplicamos nas linhas de M as operações "Substituir R_2 por $-2R_1 + R_2$" e "Substituir R_3 por $3R_1 + R_3$" e, depois, aplicamos nas colunas de M as operações correspondentes "Substituir C_2 por $-2C_1 + C_2$" e "Substituir C_3 por $3C_1 + C_3$", obtendo

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right] \quad \text{e, depois,} \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right]$$

Em seguida, aplicamos nas linhas de M a operação "Substituir R_3 por $-2R_2 + R_3$" e, depois, aplicamos nas colunas de M a operação correspondente "Substituir C_3 por $-2C_2 + C_3$", obtendo

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right] \quad \text{e, depois,} \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right]$$

Assim, A foi diagonalizada. Escrevemos

$$P = \begin{bmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e, portanto,} \quad D = P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

Enfatizamos que P é a transposta da metade à direita da matriz final.

Formas quadráticas

Começamos com uma definição.

DEFINIÇÃO A Dizemos que uma aplicação $q: V \to K$ é uma *forma quadrática* se $q(v) = f(v, v)$, para alguma forma bilinear simétrica f de V.

Se $1 + 1 \neq 0$ em K, então a forma bilinear f pode ser obtida de volta da forma quadrática q usando a *forma polar* de f a seguir.

$$f(u, v) = \tfrac{1}{2}[q(u + v) - q(u) - q(v)]$$

Seja, agora, f representada por uma matriz simétrica $A = [a_{ij}]$ e suponha que $1 + 1 \neq 0$. Escrevendo $X = [x_i]$ para um vetor coluna de variáveis, q pode ser representada na forma

$$q(X) = f(X, X) = X^T AX = \sum_{i,j} a_{ij}x_i x_j = \sum_i a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_i x_j$$

Essa expressão formal nas variáveis x_i também é denominada *forma quadrática*.

A saber, temos a definição alternativa seguinte.

DEFINIÇÃO B Uma *forma quadrática* nas variáveis $x_1, x_2, ..., x_n$ é um polinômio tal que cada um de seus termos tem grau dois, isto é,

$$q(x_1, x_2, \ldots, x_n) = \sum_i c_i x_i^2 + \sum_{i<j} d_{ij}x_i x_j$$

Usando $1 + 1 \neq 0$, a forma quadrática q da Definição B determina uma matriz simétrica $A = [a_{ij}]$, em que $a_{ii} = c_i$ e $a_{ij} = a_{ji} = \frac{1}{2}d_{ij}$. Assim, as Definições A e B são, essencialmente, iguais.

Se a representação matricial A de q é diagonal, então q tem a representação diagonal

$$q(X) = X^T AX = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2$$

Assim, o polinômio quadrático que representa q não tem termos mistos, com duas variáveis distintas. Além disso, pelo Teorema 12.4, toda forma quadrática tem alguma representação desse tipo (se $1 + 1 \neq 0$).

12.6 FORMAS BILINEARES SIMÉTRICAS REAIS, LEI DA INÉRCIA

Nesta seção tratamos de formas bilineares simétricas e formas quadráticas em espaços vetoriais V sobre o corpo dos reais $\mathbf{R}$. A natureza especial de $\mathbf{R}$ enseja uma teoria independente. O teorema seguinte é o resultado principal (demonstrado no Problema 12.14).

Teorema 12.5 Seja f uma forma bilinear simétrica de V sobre $\mathbf{R}$. Então existe uma base de V na qual f é representada por uma matriz diagonal. Qualquer outra representação matricial diagonal de f tem o mesmo número $\boldsymbol{p}$ de entradas positivas e o mesmo número $\boldsymbol{n}$ de entradas negativas.

Esse resultado é, às vezes, denominado *Lei da Inércia* ou *Teorema de Sylvester*. O *posto* e a *assinatura* de uma forma bilinear simétrica f são denotados e definidos por

$$\text{pos}(f) = \boldsymbol{p} + \boldsymbol{n} \text{ e } \text{sig}(f) = \boldsymbol{p} - \boldsymbol{n}$$

Pelo Teorema 12.5, esses valores estão determinados de modo único.

Dizemos que uma forma bilinear simétrica f é

(i) *positiva* se $q(v) = f(v, v) > 0$, para cada $v \neq 0$;
(ii) *não negativa* se $q(v) = f(v, v) \geq 0$, para cada v.

Exemplo 12.3 Seja f o produto escalar de $\mathbf{R}^n$. Vimos que f é uma forma bilinear simétrica de $\mathbf{R}^n$. Observamos que f também é positiva. De fato, dado qualquer $u = (a_i) \neq 0$ de $\mathbf{R}^n$,

$$f(u,u) = a_1^2 + a_2^2 + \cdots + a_n^2 > 0$$

Nas Seções 9.6 e 12.5 vimos como diagonalizar uma forma quadrática real q ou, equivalentemente, uma matriz simétrica real A, por meio de uma matriz de transição ortogonal P. Se P for, tão somente, não singular, podemos representar q na forma diagonal com todas entradas não nulas iguais a 1 e a –1. De fato, temos o corolário seguinte.

Corolário 12.6 Qualquer forma quadrática real q tem uma única representação do tipo

$$q(x_1, x_2, \ldots, x_n) = x_1^2 + \cdots + x_{\boldsymbol{p}}^2 - x_{\boldsymbol{p}+1}^2 - \cdots - x_r^2$$

em que $r = \boldsymbol{p} + \boldsymbol{n}$ é o posto da forma q.

Corolário 12.6 (Forma Alternativa) Qualquer matriz simétrica real A é congruente a uma única matriz diagonal

$$D = \text{diag}(I_{\boldsymbol{p}}, -I_{\boldsymbol{n}}, 0)$$

em que $r = \boldsymbol{p} + \boldsymbol{n}$ é o posto de A.

12.7 FORMAS HERMITIANAS

Seja V um espaço vetorial de dimensão finita sobre o corpo dos complexos $\mathbf{C}$. Uma *forma hermitiana* de V é uma aplicação $f: V \times V \to \mathbf{C}$ tal que, dados $a, b \in \mathbf{C}$ e $u_i, v \in V$ quaisquer

(i) $f(au_1 + bu_2,\ \ v) = af(u_1, v) + bf(u_2, v)$,
(ii) $f(u, v) = \overline{f(v, u)}$.

(Como sempre, $\bar{k}$ denota o complexo conjugado de $k \in \mathbf{C}$.)

Usando (i) e (ii), obtemos

$$\begin{aligned} f(u,\ \ av_1 + bv_2) &= \overline{f(av_1 + bv_2,\ \ u)} = \overline{af(v_1, u) + bf(v_2, u)} \\ &= \overline{af(v_1, u)} + \overline{bf(v_2, u)} = \bar{a}f(u, v_1) + \bar{b}f(u, v_2) \end{aligned}$$

Assim,

(iii) $f(u,\ \ av_1 + bv_2) = \bar{a}f(u, v_1) + \bar{b}f(u, v_2)$.

Como antes, a condição (i) significa que f é *linear na primeira variável*. Por outro lado, a condição (iii) significa que f é *linear conjugada na segunda variável*. Além disso, a condição (ii) acarreta $f(v, v) = \overline{f(v, v)}$ e, portanto, $f(v, v)$ é real, para cada $v \in V$.

Os resultados das Seções 12.5 e 12.6 para formas simétricas têm seus análogos para formas hermitianas. Assim, a aplicação $q: V \to \mathbf{C}$, definida por $q(v) = f(v, v)$ é denominada *forma quadrática hermitiana* ou *forma quadrática complexa* associada à forma hermitiana f. Podemos obter f de volta de q usando a forma polar

$$f(u, v) = \tfrac{1}{4}[q(u + v) - q(u - v)] + \tfrac{i}{4}[q(u + iv) - q(u - iv)]$$

Seja, agora, $S = \{u_1, ..., u_n\}$ uma base de V. A matriz $H = [h_{ij}]$ dada por $h_{ij} = f(u_i, u_j)$ é denominada *representação matricial* de f na base S. Por (ii), $f(u_i, u_j) = \overline{f(u_j, u_i)}$, portanto, H é hermitiana e, em particular, as entradas diagonais de H são reais. Assim, qualquer representação matricial diagonal de f tem apenas entradas reais.

O teorema seguinte (demonstrado no Problema 12.47) é o análogo complexo do Teorema 12.5 relativo a formas bilineares simétricas reais.

Teorema 12.7 Seja f uma forma hermitiana de V sobre $\mathbf{C}$. Então existe uma base de V na qual f é representada por uma matriz diagonal. Qualquer outra representação matricial diagonal de f tem o mesmo número $\boldsymbol{p}$ de entradas positivas e o mesmo número $\boldsymbol{n}$ de entradas negativas.

Novamente, o *posto* e a *assinatura* de uma forma hermitiana f são denotados e definidos por

$$\text{pos}(f) = \boldsymbol{p} + \boldsymbol{n} \text{ e } \text{sig}(f) = \boldsymbol{p} - \boldsymbol{n}$$

Pelo Teorema 12.7, esses valores estão determinados de modo único.

Analogamente, dizemos que uma forma hermitiana f é

(i) *positiva* se $q(v) = f(v, v) > 0$, para cada $v \neq 0$;
(ii) *não negativa* se $q(v) = f(v, v) \geq 0$, para cada v.

Exemplo 12.4 Seja f o produto escalar de $\mathbf{C}^n$, ou seja, para quaisquer $u = (z_i)$ e $v = (w_i)$ de $\mathbf{C}^n$,

$$f(u, v) = u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

Então, f é uma forma hermitiana de $\mathbf{C}^n$. Além disso, f também é positiva, pois, dado qualquer $u = (z_i) \neq 0$ de $\mathbf{C}^n$,

$$f(u, u) = z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 > 0$$

Problemas Resolvidos

Formas bilineares

12.1 Sejam $u = (x_1, x_2, x_3)$ e $v = (y_1, y_2, y_3)$. Escreva f em notação matricial, sendo

$$f(u, v) = 3x_1y_1 - 2x_1y_3 + 5x_2y_1 + 7x_2y_2 - 8x_2y_3 + 4x_3y_2 - 6x_3y_3$$

Seja $A = [a_{ij}]$, em que a_{ij} é o coeficiente de x_iy_j. Então

$$f(u, v) = X^TAY = [x_1, x_2, x_3]\begin{bmatrix} 3 & 0 & -2 \\ 5 & 7 & -8 \\ 0 & 4 & -6 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

12.2 Seja A uma matriz $n \times n$ sobre K. Mostre que a aplicação f definida por $f(X, Y) = X^TAY$ é uma forma bilinear de K^n.

Dados $a, b \in K$ e $X_i, Y_i \in K^n$ quaisquer,

$$\begin{aligned} f(aX_1 + bX_2,\ Y) &= (aX_1 + bX_2)^TAY = (aX_1^T + bX_2^T)AY \\ &= aX_1^TAY + bX_2^TAY = af(X_1, Y) + bf(X_2, Y) \end{aligned}$$

Logo, f é linear na primeira variável. Também,

$$f(X,\ aY_1 + bY_2) = X^TA(aY_1 + bY_2) = aX^TAY_1 + bX^TAY_2 = af(X, Y_1) + bf(X, Y_2)$$

Logo, f é linear na segunda variável e, portanto, é uma forma bilinear de K^n.

12.3 Seja f a forma bilinear de $\mathbf{R}^2$ definida por

$$f[(x_1, x_2), \quad (y_1, y_2)] = 2x_1y_1 - 3x_1y_2 + 4x_2y_2$$

(a) Encontre a matriz A de f na base $\{u_1 = (1, 0), u_2 = (1, 1)\}$.

(b) Encontre a matriz B de f na base $\{v_1 = (2, 1), v_2 = (1, -1)\}$.

(c) Encontre a matriz de mudança de base P da base $\{u_i\}$ para a base $\{v_i\}$ e verifique que $B = P^TAP$.

(a) Tomando $A = [a_{ij}]$, com $a_{ij} = f(u_i, u_j)$, obtemos

$$a_{11} = f[(1,0), \quad (1,0)] = 2 - 0 - 0 = 2, \qquad a_{21} = f[(1,1), \quad (1,0)] = 2 - 0 + 0 = 2$$
$$a_{12} = f[(1,0), \quad (1,1)] = 2 - 3 - 0 = -1, \qquad a_{22} = f[(1,1), \quad (1,1)] = 2 - 3 + 4 = 3$$

Assim, $A = \begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$ é a matriz de f na base $\{u_1, u_2\}$.

(b) Tomando $B = [b_{ij}]$, com $b_{ij} = f(v_i, v_j)$, obtemos

$$b_{11} = f[(2,1), \quad (2,1)] = 8 - 6 + 4 = 6, \qquad b_{21} = f[(1,-1), \quad (2,1)] = 4 - 3 - 4 = -3$$
$$b_{12} = f[(2,1), \quad (1,-1)] = 4 + 6 - 4 = 6, \qquad b_{22} = f[(1,-1), \quad (1,-1)] = 2 + 3 + 4 = 9$$

Assim, $B = \begin{bmatrix} 6 & 6 \\ -3 & 9 \end{bmatrix}$ é a matriz de f na base $\{v_1, v_2\}$.

(c) Escrevemos v_1 e v_2 em termos de u_1 e u_2, ou seja, $v_1 = u_1 + u_2$ e $v_2 = 2u_1 - u_2$. Então

$$P = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \qquad P^T = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

e

$$P^TAP = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ -3 & 9 \end{bmatrix} = B$$

12.4 Demonstre o Teorema 12.1. Seja V um espaço vetorial de dimensão n sobre K. Seja $\{\phi_1, \ldots, \phi_n\}$ uma base qualquer do espaço dual V^*. Então $\{f_{ij} : i, j = 1, ..., n\}$ é uma base de $B(V)$, em que f_{ij} é definida por $f_{ij}(u, v) = \phi_i(u)\phi_j(v)$. Assim, dim $B(V) = n^2$.

Seja $\{u_1, ..., u_n\}$ a base de V dual de $\{\phi_i\}$. Mostremos que $\{f_{ij}\}$ gera $B(V)$. Seja $f \in B(V)$ dado e denotemos $f(u_i, u_j) = a_{ij}$. Afirmamos que $f = \sum_{i,j} a_{ij} f_{ij}$. Basta mostrar que

$$f(u_s, u_t) = \left(\sum a_{ij} f_{ij}\right)(u_s, u_t) \qquad \text{com} \qquad s, t = 1, ..., n$$

Temos

$$\left(\sum a_{ij} f_{ij}\right)(u_s, u_t) = \sum a_{ij} f_{ij}(u_s, u_t) = \sum a_{ij}\phi_i(u_s)\phi_j(u_t) = \sum a_{ij}\delta_{is}\delta_{jt} = a_{st} = f(u_s, u_t)$$

conforme afirmado. Logo, $\{f_{ij}\}$ gera $B(V)$. Agora suponha que $\sum a_{ij}f_{ij} = 0$. Então, para $s, t = 1, ..., n$ quaisquer,

$$0 = 0(u_s, u_t) = \left(\sum a_{ij} f_{ij}\right)(u_s, u_t) = a_{rs}$$

A última igualdade segue como na primeira parte. Assim, $\{f_{ij}\}$ é independente, portanto, uma base de $B(V)$.

12.5 Demonstre o Teorema 12.2. Seja P a matriz de mudança de base de uma base S para uma base S'. Se A é a representação matricial da forma bilinear f em relação à base S, então $B = P^TAP$ é a representação matricial de f em relação à base S'.

Sejam $u, v \in V$. Como P é a matriz de mudança de base de S para S', temos $P[u]_{S'} = [u]_S$ e, também, $P[v]_{S'} = [v]_S$, de modo que $[u]_S^T = [u]_{S'}^T P^T$. Assim,

$$f(u, v) = [u]_S^T A[v]_S = [u]_{S'}^T P^T AP[v]_{S'}$$

Como u e v são elementos arbitrários de V, concluímos que P^TAP é a matriz de f na base S'.

Formas bilineares simétricas, formas quadráticas

12.6 Encontre a matriz simétrica que corresponde à forma quadrática dada.

(a) $q(x,y,z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$,

(b) $q'(x,y,z) = 3x^2 + xz - 2yz$, (c) $q''(x,y,z) = 2x^2 - 5y^2 - 7z^2$

A matriz simétrica $A = [a_{ij}]$ que representa $q(x_1, ..., x_n)$ tem entradas diagonais a_{ii} iguais aos coeficientes do termo ao quadrado x_i^2 e entradas não diagonais a_{ij} e a_{ji} iguais à metade do coeficiente do termo misto x_ix_j. Assim,

$$\text{(a)}\quad A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & -1 & -3 \\ 4 & -3 & 1 \end{bmatrix}, \text{(b)}\quad A' = \begin{bmatrix} 3 & 0 & \frac{1}{2} \\ 0 & 0 & -1 \\ \frac{1}{2} & -1 & 0 \end{bmatrix}, \text{(c)}\quad A'' = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -7 \end{bmatrix}$$

A terceira matriz A'' é diagonal porque a forma quadrática q'' é diagonal, isto é, não tem termos mistos.

12.6 Encontre a forma quadrática $q(X)$ que corresponde à matriz simétrica dada.

$$\text{(a)}\quad A = \begin{bmatrix} 5 & -3 \\ -3 & 8 \end{bmatrix}, \text{(b)}\quad B = \begin{bmatrix} 4 & -5 & 7 \\ -5 & -6 & 8 \\ 7 & 8 & -9 \end{bmatrix}, \text{(c)}\quad C = \begin{bmatrix} 2 & 4 & -1 & 5 \\ 4 & -7 & -6 & 8 \\ -1 & -6 & 3 & 9 \\ 5 & 8 & 9 & 1 \end{bmatrix}$$

A forma quadrática $q(X)$ que corresponde a uma matriz simétrica M é definida por $q(X) = X^TMX$, em que $X = [x_i]$ é o vetor coluna de variáveis.

(a) Calculando,

$$q(x,y) = X^TAX = [x,y]\begin{bmatrix} 5 & -3 \\ -3 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [5x - 3y, \quad -3x + 8y]\begin{bmatrix} x \\ y \end{bmatrix}$$
$$= 5x^2 - 3xy - 3xy + 8y^2 = 5x^2 - 6xy + 8y^2$$

Conforme esperado, o coeficiente 5 do termo ao quadrado x^2 e o coeficiente 8 do termo ao quadrado y^2 são os elementos diagonais de A. O coeficiente –6 do termo misto xy é a soma dos elementos não diagonais –3 e –3 de A (ou, o dobro do elemento não diagonal –3, já que A é simétrica).

(b) Como B é quadrada de ordem 3, temos três variáveis, digamos, x, y, z ou x_1, x_2, x_3. Então,

$$q(x,y,z) = 4x^2 - 10xy - 6y^2 + 14xz + 16yz - 9z^2$$

ou

$$q(x_1,x_2,x_3) = 4x_1^2 - 10x_1x_2 - 6x_2^2 + 14x_1x_3 + 16x_2x_3 - 9x_3^2$$

Aqui usamos que os coeficientes dos termos ao quadrado x_1^2, x_2^2, x_3^2 (ou x^2, y^2, z^2) são os respectivos elementos diagonais 4, –6, –9 de B e que os coeficientes dos termos mistos x_ix_j são a soma dos elementos não diagonais b_{ij} e b_{ji} (ou duas vezes b_{ij}, já que $b_{ij} = b_{ji}$).

(c) Como C é quadrada de ordem 4, temos quatro variáveis, digamos,

$$q(x_1,x_2,x_3,x_4) = 2x_1^2 - 7x_2^2 + 3x_3^2 + x_4^2 + 8x_1x_2 - 2x_1x_3$$
$$+ 10x_1x_4 - 12x_2x_3 + 16x_2x_4 + 18x_3x_4$$

12.8 Seja $A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{bmatrix}$. Aplique o Algoritmo 12.1 para encontrar uma matriz não singular P tal que $D = P^TAP$ seja diagonal e encontre a assinatura sig(A) de A.

Começamos formando a matriz $M = [A, I]$, como segue.

$$M = [A, I] = \left[\begin{array}{rrr|rrr} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right]$$

Usando a_{11} como pivô, aplicamos as operações "Substituir R_2 por $3R_1 + R_2$" e "Substituir R_3 por $-2R_1 + R_3$" com as linhas de M e depois aplicamos as operações "Substituir C_2 por $3C_1 + C_2$" e "Substituir C_3 por $-2C_1 + C_3$" com as colunas de A, obtendo

$$\left[\begin{array}{rrr|rrr} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right] \quad \text{e, então,} \quad \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right].$$

Agora aplicamos a operação "Substituir R_3 por $R_2 + 2R_3$" e depois aplicamos a operação "Substituir C_3 por $C_2 + 2C_3$", obtendo

$$\left[\begin{array}{rrrrrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right] \quad \text{e, então,} \quad \left[\begin{array}{rrrrrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right]$$

Agora, A foi diagonalizada e a transposta de P está na metade à direita de M. Assim, tomamos

$$P = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{e, então,} \quad D = P^TAP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{bmatrix}$$

Observe que D tem $\boldsymbol{p} = 2$ elementos diagonais positivos e $\boldsymbol{n} = 1$ entrada negativa. Assim, a assinatura de A é $\text{sig}(A) = \boldsymbol{p} - \boldsymbol{n} = 2 - 1 = 1$.

12.9 Justifique o Algoritmo 12.1 que diagonaliza (via congruência) uma matriz simétrica A.

Considere a matriz em blocos $M = [A, I]$. O algoritmo aplica uma sequência de operações elementares com as linhas, seguida pela sequência correspondente de operações com as colunas do lado esquerdo de M, que é a matriz A. Isso equivale a pré-multiplicar A por uma sequência de matrizes elementares, digamos, $E_1, E_2, \ldots, E_r$, e pós-multiplicar A pelas transpostas de E_i. Assim, quando o algoritmo termina, a matriz diagonal D à esquerda de M é igual a

$$D = E_r \cdots E_2 E_1 A E_1^T E_2^T \cdots E_r^T = QAQ^T, \qquad \text{com} \qquad Q = E_r \cdots E_2 E_1$$

Por outro lado, o algoritmo somente aplica as operações elementares com as linhas à matriz identidade I do lado direito de M. Assim, quando o algoritmo termina, a matriz à direita de M é igual a

$$E_r \cdots E_2 E_1 I = E_r \cdots E_2 E_1 = Q$$

Tomando $P = Q^T$, obtemos $D = P^TAP$, que é a diagonalização de A via congruência.

12.10 Demonstre o Teorema 12.4. Seja f uma forma bilinear simétrica de V sobre K (em que $1 + 1 \neq 0$). Então V tem uma base na qual f é representada por uma matriz diagonal.

O Algoritmo 12.1 mostra que cada matriz simétrica sobre K é congruente a uma matriz diagonal. Essa afirmação é equivalente à afirmação de que f tem uma representação diagonal.

12.11 Seja q a forma quadrática associada à forma bilinear simétrica f. (Suponha que $1 + 1 \neq 0$).Verifique a identidade polar $f(u, v) = \frac{1}{2}[q(u + v) - q(u) - q(v)]$.

Temos

$$\begin{aligned} q(u+v) - q(u) - q(v) &= f(u+v, \; u+v) - f(u,u) - f(v,v) \\ &= f(u,u) + f(u,v) + f(v,u) + f(v,v) - f(u,u) - f(v,v) = 2f(u,v) \end{aligned}$$

Se $1 + 1 \neq 0$, podemos dividir tudo por 2 para obter a identidade polar.

12.12 Considere a forma quadrática $q(x, y) = 3x^2 + 2xy - y^2$ e a substituição linear

$$x = s - 3t, \qquad y = 2s + t$$

(a) Reescreva $q(x, y)$ em notação matricial e encontre a matriz A que representa $q(x, y)$.

(b) Reescreva a substituição linear usando notação matricial e encontre a matriz P correspondente à substituição.

(c) Encontre $q(s, t)$ usando substituição direta.

(d) Encontre $q(s, t)$ usando notação matricial.

(a) Temos $q(x, y) = [x, y]\begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Assim, $A = \begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}$ e $q(X) = X^TAX$, com $X = [x, y]^T$.

(b) Temos $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix}$. Assim, $P = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}$, com $X = \begin{bmatrix} x \\ y \end{bmatrix}$, $Y = \begin{bmatrix} s \\ t \end{bmatrix}$ e $X = PY$.

(c) Substituímos x e y em q para obter

$$\begin{aligned} q(s, t) &= 3(s - 3t)^2 + 2(s - 3t)(2s + t) - (2s + t)^2 \\ &= 3(s^2 - 6st + 9t^2) + 2(2s^2 - 5st - 3t^2) - (4s^2 + 4st + t^2) = 3s^2 - 32st + 20t^2 \end{aligned}$$

(d) Temos $q(X) = X^TAX$ e $X = PY$. Assim, $X^T = Y^TP^T$ e, portanto,

$$\begin{aligned} q(s, t) = q(Y) = Y^TP^TAPY &= [s, t]\begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix} \\ &= [s, t]\begin{bmatrix} 3 & -16 \\ -16 & 20 \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix} = 3s^2 - 32st + 20t^2 \end{aligned}$$

[Conforme esperado, são iguais os resultados de (c) e (d).]

12.13 Considere uma matriz diagonal $A = \text{diag}(a_1, ..., a_n)$ qualquer sobre K. Mostre que, dados quaisquer escalares não nulos $k_1, \ldots, k_n \in K$, A é congruente à matriz diagonal D de entradas $a_1k_1^2, \ldots, a_nk_n^2$. Além disso, demonstre as afirmações dadas.

(a) Se $K = \mathbf{C}$, podemos escolher D tal que cada entrada diagonal seja 1 ou 0.

(b) Se $K = \mathbf{R}$, podemos escolher D tal que cada entrada diagonal seja 1, –1 ou 0.

Seja $P = \text{diag}(k_1, ..., k_n)$. Então,

$$D = P^TAP = \text{diag}(k_i)\ \text{diag}(a_i)\ \text{diag}(k_i) = \text{diag}(a_1k_1^2, \ldots, a_nk_n^2)$$

(a) Seja $P = \text{diag}(b_i)$, com $b_i = \begin{cases} 1/\sqrt{a_i} & \text{se } a_i \neq 0 \\ 1 & \text{se } a_i = 0 \end{cases}$

Então todas as entradas diagonais de P^TAP são iguais a 1 ou 0.

(b) Seja $P = \text{diag}(b_i)$, com $b_i = \begin{cases} 1/\sqrt{|a_i|} & \text{se } a_i \neq 0 \\ 1 & \text{se } a_i = 0 \end{cases}$

Então todas as entradas diagonais de P^TAP são iguais a 1, –1 ou 0.

OBSERVAÇÃO Enfatizamos que (b) não segue válido trocando "congruência" por "congruência hermitiana".

12.14 Demonstre o Teorema 12.5. Seja f uma forma bilinear simétrica de V sobre $\mathbf{R}$. Então existe uma base de V na qual f é representada por uma matriz diagonal. Qualquer outra representação matricial diagonal de f tem o mesmo número $\boldsymbol{p}$ de entradas positivas e o mesmo número $\boldsymbol{n}$ de entradas negativas.

Pelo Teorema 12.4, existe uma base $\{u_1, ..., u_n\}$ de V na qual f é representada por uma matriz diagonal com, digamos, $\boldsymbol{p}$ entradas positivas e $\boldsymbol{n}$ negativas. Agora, seja $\{w_1, ..., w_n\}$ uma outra base de V, na qual f é representada por uma matriz diagonal com, digamos, $\boldsymbol{p}'$ entradas positivas e $\boldsymbol{n}'$ negativas. Podemos supor, sem perda de generalidade, que as entradas positivas aparecem antes das demais, em cada matriz. Como $\text{pos}(f) = \boldsymbol{p} + \boldsymbol{n} = \boldsymbol{p}' + \boldsymbol{n}'$, basta provar que $\boldsymbol{p} = \boldsymbol{p}'$.

Sejam U o espaço gerado por $u_1, ..., u_p$ e W o espaço gerado por $w_{p'+1}, \ldots, w_n$. Então $f(v, v) > 0$ para cada vetor não nulo $v \in U$, e $f(v, v) \leq 0$ para cada vetor não nulo $v \in W$. Logo, $U \cap W = \{0\}$. Observe que dim $U = \boldsymbol{p}$ e dim $W = n - \boldsymbol{p}'$. Assim,

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \boldsymbol{p} + (n - \boldsymbol{p}') - 0 = \boldsymbol{p} - \boldsymbol{p}' + n$$

Como $\dim(U + W) \leq \dim V = n$, resulta $\boldsymbol{p} - \boldsymbol{p}' + n \leq n$ ou $\boldsymbol{p} \leq \boldsymbol{p}'$. Analogamente, mostramos que $\boldsymbol{p}' \leq \boldsymbol{p}$, de modo que $\boldsymbol{p} = \boldsymbol{p}'$, como queríamos mostrar.

OBSERVAÇÃO A validade do Teorema 12.5 e a demonstração apresentada dependem tão somente do conceito de positividade. Assim, o teorema também é válido em qualquer subcorpo K do corpo real $\mathbf{R}$, como o corpo racional $\mathbf{Q}$.

Formas quadráticas reais positivas

12.15 Prove que são equivalentes as definições dadas de positividade.

(a) São positivas as entradas diagonais de qualquer representação diagonal de q.

(b) $q(Y) > 0$, para qualquer vetor não nulo Y de $\mathbf{R}^n$.

Digamos que $q(Y) = a_1y_1^2 + a_2y_2^2 + \cdots + a_ny_n^2$. Se todos os coeficientes forem positivos, claramente $q(Y) > 0$, sempre que $Y \neq 0$. Assim, (a) implica (b). Reciprocamente, suponha que (a) seja falso, ou seja, suponha que alguma entrada diagonal $a_k \leq 0$. Seja $e_k = (0, ..., 1, ..., 0)$ o vetor com todas as entradas nulas, exceto por um 1 na k-ésima posição. Então $q(e_k) = a_k$ não é positivo, de modo que (b) é falso. Assim, (b) implica (a). Em vista disso, (a) e (b) são equivalentes.

12.16 Decida se a forma quadrática dada é positiva ou não.

(a) $q(x, y, z) = x^2 + 2y^2 - 4xz - 4yz + 7z^2$

(b) $q(x, y, z) = x^2 + y^2 + 2xz + 4yz + 3z^2$

Diagonalizamos a matriz simétrica A correspondente a q via congruência.

(a) Aplicamos as operações "Substituir R_3 por $2R_1 + R_3$" e "Substituir C_3 por $2C_1 + C_3$" e, depois, "Substituir R_3 por $R_2 + R_3$" e "Substituir C_3 por $C_2 + C_3$", obtendo

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 7 \end{bmatrix} \simeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 3 \end{bmatrix} \simeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A representação diagonal de q contém as entradas positivas 1, 2, 1 na diagonal. Assim, q é positiva.

(b) Temos

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} \simeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 2 \end{bmatrix} \simeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Há uma entrada negativa -2 na representação diagonal de q. Assim, q não é positiva.

12.17 Mostre que $q(x, y) = ax^2 + bxy + cy^2$ é positiva se, e só se, $a > 0$ e o discriminante $D = b^2 - 4ac < 0$.

Suponha que $v = (x, y) \neq 0$. Então $x \neq 0$ ou $y \neq 0$; digamos, $y \neq 0$. Seja $t = x/y$. Então

$$q(v) = y^2[a(x/y)^2 + b(x/y) + c] = y^2(at^2 + bt + c)$$

Contudo, as afirmações seguintes são equivalentes.

(i) $s = at^2 + bt + c$ é positiva para cada valor de t.

(ii) $s = at^2 + bt + c$ se situa acima do eixo t.

(iii) $a > 0$ e $D = b^2 - 4ac < 0$.

Assim, q é positiva se, e só se, $a > 0$ e $D < 0$. [*Observação*: $D < 0$ é o mesmo que $\det(A) > 0$, sendo A a matriz simétrica correspondente a q.]

12.18 Decida se a forma quadrática dada é positiva ou não.

(a) $q(x,y) = x^2 - 4xy + 7y^2$, (b) $q(x,y) = x^2 + 8xy + 5y^2$, (c) $q(x,y) = 3x^2 + 2xy + y^2$

Calculamos o discriminante $D = b^2 - 4ac$ e usamos o Problema 12.17.

(a) $D = 16 - 28 = -12$. Como $a = 1 > 0$ e $D < 0$, q é positiva.

(b) $D = 64 - 20 = 44$. Como $D > 0$, q não é positiva.

(c) $D = 4 - 12 = -8$. Como $a = 3 > 0$ e $D < 0$, q é positiva.

Formas hermitianas

12.19 Decida se a matriz dada é hermitiana.

(a) $\begin{bmatrix} 2 & 2+3i & 4-5i \\ 2-3i & 5 & 6+2i \\ 4+5i & 6-2i & -7 \end{bmatrix}$, (b) $\begin{bmatrix} 3 & 2-i & 4+i \\ 2-i & 6 & i \\ 4+i & i & 7 \end{bmatrix}$, (c) $\begin{bmatrix} 4 & -3 & 5 \\ -3 & 2 & 1 \\ 5 & 1 & -6 \end{bmatrix}$

Uma matriz complexa $A = [a_{ij}]$ é hermitiana se $A^* = A$, ou seja, se $a_{ij} = \bar{a}_{ji}$.

(a) É hermitiana, por ser igual à transposta conjugada.

(b) Não é hermitiana, mesmo sendo simétrica; por exemplo, o conjugado de i é $-i$.

(c) É hermitiana. Observe que uma matriz real é hermitiana se, e só se, é simétrica.

12.20 Seja A uma matriz hermitiana. Mostre que $f(X, Y) = X^T A \bar{Y}$ define uma forma hermitiana de $\mathbf{C}^n$.

Dados $a, b \in \mathbf{C}$ e $X_1, X_2, Y \in \mathbf{C}^n$ quaisquer,

$$\begin{aligned} f(aX_1 + bX_2,\ Y) &= (aX_1 + bX_2)^T A\bar{Y} = (aX_1^T + bX_2^T)A\bar{Y} \\ &= aX_1^T A\bar{Y} + bX_2^T A\bar{Y} = af(X_1, Y) + bf(X_2, Y) \end{aligned}$$

Logo, f é linear na primeira variável. Também

$$\overline{f(X,Y)} = \overline{X^T A\bar{Y}} = \overline{(X^T A\bar{Y})^T} = \overline{\bar{Y}^T A^T X} = Y^T A^* \bar{X} = Y^T A\bar{X} = f(Y,X)$$

Assim, f é uma forma hermitiana de $\mathbf{C}^n$.

OBSERVAÇÃO Utilizamos o fato de que $X^T A \bar{Y}$ é escalar e, portanto, igual à transposta.

12.21 Sejam f uma forma hermitiana de V e H a matriz de f na base $S = \{u_i\}$ de V. Prove as afirmações dadas.

(a) $f(u, v) = [u]_S^T H \overline{[v]}_S$, para quaisquer $u, v \in V$.

(b) Se P é a matriz de mudança de base de S para uma nova base S' de V, então $B = P^T H \bar{P}$ (ou $B = Q^* HQ$, em que $Q = \bar{P}$) é a matriz de f na nova base S'.

Observe que (b) é o análogo complexo do Teorema 12.2.

(a) Sejam $u, v \in V$ e suponha que $u = a_1u_1 + \cdots + a_nu_n$ e $v = b_1u_1 + \cdots + b_nu_n$. Calculando,

$$\begin{aligned} f(u, v) &= f(a_1u_1 + \cdots + a_nu_n,\ b_1u_1 + \cdots + b_nu_n) \\ &= \sum_{i,j} a_i\bar{b}_j f(u_i, v_j) = [a_1, \ldots, a_n]H[\bar{b}_1, \ldots, \bar{b}_n]^T = [u]_S^T H\overline{[v]}_S \end{aligned}$$

(b) Sendo P a matriz de mudança de base de S para S', temos $P[u]_{S'} = [u]_S$ e $P[v]_{S'} = [v]_S$, portanto, $[u]_S^T = [u]_{S'}^T P^T$ e $\overline{[v]_S} = \bar{P}\overline{[v]_{S'}}$. Assim, por (a),

$$f(u, v) = [u]_S^T H\overline{[v]_S} = [u]_{S'}^T P^T H\bar{P}\overline{[v]_{S'}}$$

Como u e v são elementos arbitrários de V, concluímos que $P^T H\bar{P}$ é a matriz de f na base S'.

12.22 Considere a matriz hermitiana $H = \begin{bmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{bmatrix}$.

Encontre uma matriz não singular P tal que $D = P^T H\bar{P}$ seja diagonal. Também encontre a assinatura de H.

Usamos um Algoritmo 12.1 modificado, que aplica as mesmas operações com as linhas, mas depois, o conjugado das operações correspondentes nas colunas. Assim, começamos formando a matriz em blocos $M = [H, I]$, como segue.

$$M = \begin{bmatrix} 1 & 1+i & 2i & 1 & 0 & 0 \\ 1-i & 4 & 2-3i & 0 & 1 & 0 \\ -2i & 2+3i & 7 & 0 & 0 & 1 \end{bmatrix}$$

Aplicamos as operações "Substituir R_2 por $(-1 + i)R_1 + R_2$" e "Substituir R_3 por $2iR_1 + R_3$" com as linhas de M e depois as operações conjugadas correspondentes "Substituir C_2 por $(-1 - i)C_1 + C_2$" e "Substituir C_3 por $-2iC_1 + C_3$" com as colunas, obtendo

$$\begin{bmatrix} 1 & 1+i & 2i & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{bmatrix} \quad \text{e, então,} \quad \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{bmatrix}$$

Agora aplicamos a operação "Substituir R_3 por $-5iR_2 + 2R_3$" e a operação conjugada correspondente "Substituir C_3 por $5iC_2 + 2C_3$", obtendo

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 0 & -19 & 5+9i & -5i & 2 \end{bmatrix} \quad \text{e, então,} \quad \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1+i & 1 & 0 \\ 0 & 0 & -38 & 5+9i & -5i & 2 \end{bmatrix}$$

Agora, H foi diagonalizada e a transposta de P está na metade à direita de M. Assim, tomamos

$$P = \begin{bmatrix} 1 & -1+i & 5+9i \\ 0 & 1 & -5i \\ 0 & 0 & 2 \end{bmatrix} \quad \text{e, então,} \quad D = P^T H\bar{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{bmatrix}.$$

Observe que D tem $\boldsymbol{p} = 2$ elementos diagonais positivos e $\boldsymbol{n} = 1$ entrada negativa. Assim, a assinatura de H é $\text{sig}(H) = \boldsymbol{p} - \boldsymbol{n} = 2 - 1 = 1$.

Formas bilineares alternadas

12.23 Demonstre o Teorema 12.3. Seja f uma forma bilinear alternada de V. Então existe uma base de V na qual f é representada por uma matriz diagonal em blocos M com blocos do tipo $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ ou 0. O número de blocos não nulos é determinado de modo único por f [porque é igual à metade do posto de f].

Se $f = 0$, o teorema é trivialmente verdadeiro. Também, se $\dim V = 1$, então $f(r_1u, r_2u) = r_1r_2f(u, u) = 0$ e, portanto, $f = 0$. Em vista disso, podemos supor que $\dim V > 1$ e que $f \neq 0$.

Como $f \neq 0$, existem vetores (não nulos) $u_1, u_2 \in V$ tais que $f(u_1, u_2) \neq 0$. De fato, multiplicando u_1 por um fator apropriado, podemos supor que $f(u_1, u_2) = 1$ e, portanto, $f(u_2, u_1) = -1$. Segue que u_1 e u_2 são linearmente independentes, pois, se $u_2 = ru_1$, digamos, então $f(u_1, u_2) = f(u_1, ru_1) = rf(u_1, u_1) = 0$. Denotando $U = \text{ger}(u_1, u_2)$, obtemos as afirmações a seguir.

(i) A representação matricial da restrição de f a U na base $\{u_1, u_2\}$ é $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

(ii) Se $u \in U$, digamos, $u = au_1 + bu_2$, então

$$f(u, u_1) = f(au_1 + bu_2, \ u_1) = -b \qquad \text{e} \qquad f(u, u_2) = f(au_1 + bu_2, \ u_2) = a$$

Seja W o conjunto daqueles vetores $w \in V$ para os quais $f(w, u_1) = 0$ e $f(w, u_2) = 0$, ou seja,

$$W = \{w \in V : f(w, u) = 0 \text{ para todo } u \in U\}$$

Afirmamos que $V = U \oplus W$. É claro que $U \cap W = \{0\}$, de modo que basta mostrar que $V = U + W$. Seja $v \in V$ dado. Tomamos

$$u = f(v, u_2)u_1 - f(v, u_1)u_2 \qquad \text{e} \qquad w = v - u \tag{1}$$

Como u é uma combinação linear de u_1 e u_2, temos $u_2, u \in U$.
Mostremos que $w \in W$. Por (1) e (ii), $f(u, u_1) = f(v, u_1)$, portanto,

$$f(w, u_1) = f(v - u, \ u_1) = f(v, u_1) - f(u, u_1) = 0$$

Analogamente, $f(u, u_2) = f(v, u_2)$ e, portanto,

$$f(w, u_2) + f(v - u, \ u_2) = f(v, u_2) - f(u, u_2) = 0$$

Então $w \in W$ e, por (1) segue que $v = u + w$, com $u \in W$. Dessa forma, mostramos que $V = U + W$. Assim, $V = U \oplus W$.
Resta observar que a restrição de f a W é uma forma bilinear alternada de W. Por indução, existe uma base $u_3, ..., u_n$ de W na qual a representação matricial da restrição de f a W tem o formato procurado. Assim, $u_1, u_2, u_3, ..., u_n$ é uma base de V na qual a representação matricial de f tem o formato procurado.

Problemas Complementares

Formas bilineares

12.24 Sejam $u = (x_1, x_2)$ e $v = (y_1, y_2)$. Decida se a aplicação dada é uma forma bilinear de $\mathbf{R}^2$ ou não.

(a) $f(u, v) = 2x_1y_2 - 3x_2y_1$, (c) $f(u, v) = 3x_2y_2$, (e) $f(u, v) = 1$,
(b) $f(u, v) = x_1 + y_2$, (d) $f(u, v) = x_1x_2 + y_1y_2$, (f) $f(u, v) = 0$

12.25 Seja f a forma bilinear de $\mathbf{R}^2$ definida por

$$f[(x_1, x_2), \ (y_1, y_2)] = 3x_1y_1 - 2x_1y_2 + 4x_2y_1 - x_2y_2$$

(a) Encontre a matriz A de f na base $\{u_1 = (1, 1), u_2 = (1, 2)\}$.

(b) Encontre a matriz B de f na base $\{v_1 = (1, –1), v_2 = (3, 1)\}$.

(c) Encontre a matriz de mudança de base P de $\{u_i\}$ para $\{v_i\}$ e verifique que $B = P^TAP$.

12.26 Seja V o espaço vetorial das matrizes quadradas de ordem 2 sobre $\mathbf{R}$. Seja $M = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ dada e defina a aplicação $f(A, B) = \text{tr}(A^TMB)$, para $A, B \in V$, em que "tr" denota o traço. (a) Mostre que f é uma forma bilinear de V. (b) Encontre a matriz de f na base

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

12.27 Seja $B(V)$ o conjunto das formas bilineares de V sobre K. Prove as afirmações seguintes.

(a) Se $f, g \in B(V)$, então $f + g$, $kg \in B(V)$, para qualquer $k \in K$.

(b) Se ϕ e σ são funcionais lineares de V, então $f(u, v) = \phi(u)\sigma(v)$ pertence a $B(V)$.

12.28 Denotemos por $[f]$ a representação matricial de uma forma bilinear f de V em relação a uma base $\{u_1, ..., u_n\}$ de V. Mostre que a aplicação $f \mapsto [f]$ é um isomorfismo de $B(V)$ sobre o espaço vetorial das matrizes quadradas de ordem n.

12.29 Seja f uma forma bilinear de V. Dado um subconjunto S qualquer de V, sejam

$$S^{\perp} = \{v \in V : f(u, v) = 0 \text{ para cada } u \in S\} \qquad \text{e} \qquad S^{\top} = \{v \in V : f(v, u) = 0 \text{ para cada } u \in S\}$$

Mostre que (a) $S^{\top}$ e $S^{\top}$ são subespaços de V; (b) $S_1 \subseteq S_2$ implica $S_2^{\perp} \subseteq S_1^{\perp}$ e $S_2^{\top} \subseteq S_1^{\top}$;
(c) $\{0\}^{\perp} = \{0\}^{\top} = V$.

12.30 Seja f uma forma bilinear de V. Prove que pos(f) = $\dim V - \dim V^{\perp} = \dim V - \dim V^{\top}$ e que, portanto, $V^{\perp} = \dim V^{\top}$.

12.31 Seja f uma forma bilinear de V. Dado qualquer $u \in V$, defina $\hat{u}:V \to K$ e $\tilde{u}:V \to K$ por $\hat{u}(x) = f(x, u)$ e $\tilde{u}(x) = f(u, x)$. Prove as afirmações seguintes.

(a) $\hat{u}$ e $\tilde{u}$ são lineares, ou seja, $\hat{u}, \tilde{u} \in V^*$.

(b) $u \mapsto \hat{u}$ e $u \mapsto \tilde{u}$ são transformações lineares de V em V^*.

(c) $\text{pos}\,(f) = \text{pos}\,(u \mapsto \hat{u}) = \text{pos}\,(u \mapsto \tilde{u})$.

12.32 Mostre que a congruência de matrizes (denotada por $\simeq$) é uma relação de equivalência, isto é,

(a) $A \simeq A$; (ii) se $A \simeq B$, então $B \simeq A$; (iii) se $A \simeq B$ e $B \simeq C$, então $A \simeq C$.

Formas bilineares simétricas, formas quadráticas

12.33 Encontre a matriz simétrica A associada à forma quadrática dada.

(a) $q(x, y, z) - 2x^2 - 8xy + y^2 - 16xz + 14yz + 5z^2$,
(b) $q(x, y, z) = x^2 - xz + y^2$,
(c) $q(x, y, z) = xy + y^2 + 4xz + z^2$
(d) $q(x, y, z) = xy + yz$

12.34 Encontre uma matriz não singular P tal que $D = P^TAP$ seja diagonal, nos casos de A dados.

(a) $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 6 \\ 2 & 6 & 7 \end{bmatrix}$, (b) $A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & 3 \\ 1 & 3 & -2 \end{bmatrix}$, (c) $A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 2 & -1 \end{bmatrix}$

12.35 Sejam $q(x, y) = 2x^2 - 6xy - 3y^2$ e $x = s + 2t$, $y = 3s - t$.

(a) Reescreva $q(x, y)$ em notação matricial e encontre a matriz A que representa a forma quadrática.

(b) Reescreva a substituição linear usando notação matricial e encontre a matriz P correspondente à substituição.

(c) Encontre $q(s, t)$ usando (i) substituição direta, (ii) notação matricial.

12.36 Para a forma quadrática $q(x, y, z)$ dada, encontre uma substituição linear não singular escrevendo as variáveis x, y, z em termos de variáveis r, s, t tais que $q(r, s, t)$ seja diagonal.

(a) $q(x, y, z) = x^2 + 6xy + 8y^2 - 4xz + 2yz - 9z^2$,
(b) $q(x, y, z) = 2x^2 - 3y^2 + 8xz + 12yz + 25z^2$,
(c) $q(x, y, z) = x^2 + 2xy + 3y^2 + 4xz + 8yz + 6z^2$.

Em cada caso, encontre o posto e a assinatura da forma.

12.37 Dê um exemplo de uma forma quadrática $q(x, y)$ tal que $q(u) = 0$ e $q(v) = 0$, mas $q(u + v) \neq 0$.

12.38 Seja $S(V)$ o conjunto de todas as formas bilineares simétricas de V. Mostre que

(a) $S(V)$ é um subespaço de $B(V)$; (b) se $\dim V = n$, então $\dim S(V) = \frac{1}{2}n(n + 1)$.

12.39 Considere um polinômio real quadrático $q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$, com $a_{ij} = a_{ji}$.

(a) Se $a_{11} \neq 0$, mostre que a substituição

$$x_1 = y_1 - \frac{1}{a_{11}}(a_{12}y_2 + \cdots + a_{1n}y_n), \qquad x_2 = y_2, \qquad \ldots, \qquad x_n = y_n$$

fornece a equação $q(x_1, \ldots, x_n) = a_{11} y_1^2 + q'(y_2, \ldots, y_n)$, em que q' também é um polinômio quadrático.

(b) Se $a_{11} = 0$, mas, digamos, $a_{12} \neq 0$, mostre que a substituição

$$x_1 = y_1 + y_2, \qquad x_2 = y_1 - y_2, \qquad x_3 = y_3, \qquad \ldots, \qquad x_n = y_n$$

fornece a equação $q(x_1, \ldots, x_n) = \sum b_{ij} y_i y_j$, em que $b_{11} \neq 0$, o que reduz esse caso ao caso (a).

OBSERVAÇÃO Esse método de diagonalização de q é conhecido como completamento do quadrado.

Formas quadráticas positivas

12.40 Decida se a forma quadrática dada é positiva ou não.

(a) $q(x, y) = 4x^2 + 5xy + 7y^2$, (c) $q(x, y, z) = x^2 + 4xy + 5y^2 + 6xz + 2yz + 4z^2$
(b) $q(x, y) = 2x^2 - 3xy - y^2$, (d) $q(x, y, z) = x^2 + 2xy + 2y^2 + 4xz + 6yz + 7z^2$

12.41 Encontre os valores de k com os quais a forma quadrática dada é positiva.

(a) $q(x, y) = 2x^2 - 5xy + ky^2$, (b) $q(x, y) = 3x^2 - kxy + 12y^2$
(c) $q(x, y, z) = x^2 + 2xy + 2y^2 + 2xz + 6yz + kz^2$

12.42 Seja A uma matriz real simétrica positiva. Mostre que $A = P^T P$, para alguma matriz não singular P.

Formas hermitianas

12.43 Modifique o Algoritmo 12.1 de tal modo que, para uma dada matriz hermitiana H, produza uma matriz não singular P tal que $D = P^T H \bar{P}$ seja diagonal.

12.44 Encontre uma matriz não singular P tal que $D = P^T H \bar{P}$ seja diagonal, nos casos de H dados.

(a) $H = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$, (b) $H = \begin{bmatrix} 1 & 2+3i \\ 2-3i & -1 \end{bmatrix}$, (c) $H = \begin{bmatrix} 1 & i & 2+i \\ -i & 2 & 1-i \\ 2-i & 1+i & 2 \end{bmatrix}$

Em cada caso, encontre o posto e a assinatura da matriz.

12.45 Seja A uma matriz complexa não singular. Mostre que $H = A^*A$ é hermitiana e positiva.

12.46 Dizemos que B é *congruente hermitianamente* a A se existir uma matriz não singular P tal que $B = P^T A \bar{P}$ ou, equivalentemente, se existir uma matriz não singular Q tal que $B = Q^*AQ$. Mostre que a congruência hermitiana é uma relação de equivalência. (*Observação*: se $P = \bar{Q}$, então $P^T A \bar{P} = Q^*AQ$.)

12.47 Demonstre o Teorema 12.7. Seja f uma forma hermitiana de V. Então existe uma base S de V na qual f é representada por uma matriz diagonal e qualquer outra representação diagonal de f tem o mesmo número $\boldsymbol{p}$ de entradas positivas e o mesmo número $\boldsymbol{n}$ de entradas negativas.

Problemas variados

12.48 Sejam e uma operação elementar com as linhas e f^* a correspondente operação conjugada com as colunas (em que cada escalar k de e é substituído pelo conjugado $\bar{k}$ em f^*). Mostre que a matriz elementar correspondente a f^* é a transposta conjugada da matriz elementar correspondente a e.

12.49 Sejam V e W espaços vetoriais sobre K. Dizemos que uma aplicação $f: V \times W \to K$ é uma *forma bilinear* de V e W se

(i) $f(av_1 + bv_2, \ w) = af(v_1, w) + bf(v_2, w)$,

(ii) $f(v, \ aw_1 + bw_2) = af(v, w_1) + bf(v, w_2)$

para quaisquer $a, b \in K$, $v_i \in V$, $w_j \in W$. Demonstre as afirmações seguintes.

(a) O conjunto $B(V, W)$ de todas as formas bilineares de V e W é um subespaço do espaço vetorial das aplicações de $V \times W$ em K.

(b) Se $\{\phi_1, \ldots, \phi_m\}$ é uma base de V^* e $\{\sigma_1, \ldots, \sigma_n\}$ uma base de W^*, então $\{f_{ij} : i = 1, ..., m, j = 1, ..., n\}$ é uma base de $B(V, W)$, sendo f_{ij} definido por $f_{ij}(v, w) = \phi_i(v)\sigma_j(w)$. Assim, $\dim B(V, W) = \dim V \dim W$.

[Observe que, se $V = W$, então obtemos o espaço $B(V)$ estudado neste capítulo.]

12.50 Seja V um espaço vetorial sobre K. Dizemos que uma aplicação $f : \overbrace{V \times V \times \ldots \times V}^{m \text{ vezes}} \to K$ é uma *forma multilinear* (ou *m-linear*) de V se f é linear em cada variável, ou seja, para cada $i = 1, ..., m$,

$$f(\ldots, \widehat{au + bv}, \ldots) = af(\ldots, \hat{u}, \ldots) + bf(\ldots, \hat{v}, \ldots)$$

em que $\widehat{\ldots}$ denota o i-ésimo elemento, sendo todos os demais elementos mantidos fixados. Dizemos que uma forma m-linear é *alternada* se $f(v_1, ..., v_m) = 0$, sempre que $v_i = v_j$, para $i \neq j$. Demonstre as afirmações seguintes.

(a) O conjunto $B_m(V)$ de todas as formas m-lineares de V é um subespaço do espaço vetorial das aplicações de $V \times V \times \cdots \times V$ em K.

(b) O conjunto $A_m(V)$ das formas m-lineares alternadas de V é um subespaço de $B_m(V)$.

OBSERVAÇÃO 1 Se $m = 2$, obtemos o espaço $B(V)$ estudado neste capítulo.

OBSERVAÇÃO 2 Se $V = K^m$, então a função determinante é uma forma alternada m-linear de V.

Respostas dos Problemas Complementares

Notação: $M = [R_1; \quad R_2; \quad \ldots]$ denota uma matriz M de linhas $R_1, R_2, \ldots$

12.24 (a) É, (b) Não é, (c) É, (d) Não é, (e) Não é, (f) É.

12.25 (a) $A = [4, 1; \ 7, 3]$, (b) $B = [0, -4; \ 20, 32]$, (c) $P = [3, 5; \ -2, -2]$

12.26 (b) $[1, 0, 2, 0; \ 0, 1, 0, 2; \ 3, 0, 5, 0; \ 0, 3, 0, 5]$

12.33 (a) $[2, -4, -8; \ -4, 1, 7; \ -8, 7, 5]$, (b) $[1, 0, -\frac{1}{2}; \ 0, 1, 0; \ -\frac{1}{2}, 0, 0]$,
(c) $[0, \frac{1}{2}, 2; \ \frac{1}{2}, 1, 0; \ 2, 0, 1]$, (d) $[0, \frac{1}{2}, 0; \ \frac{1}{2}, 0, 1; \ \frac{1}{2}, 0, \frac{1}{2}; \ 0, \frac{1}{2}, 0]$

12.34 (a) $P = [1, 0, -2; \ 0, 1, -2; \ 0, 0, 1]$, $D = \text{diag}(1, 3, -9)$;
(b) $P = [1, 2, -11; \ 0, 1, -5; \ 0, 0, 1]$, $D = \text{diag}(1, 1, -28)$;
(c) $P = [1, 1, -1, -4; \ 0, 1, -1, -2; \ 0, 0, 1, 0; \ 0, 0, 0, 1]$, $D = \text{diag}(1, 1, 0, -9)$

12.35 $A = [2, -3; \ -3, -3]$, $P = [1, 2; \ 3, -1]$, $q(s, t) = -43s^2 - 4st + 17t^2$

12.36 (a) $x = r - 3s - 19t$, $y = s + 7t$, $z = t$; $q(r, s, t) = r^2 - s^2 + 36t^2$;
(b) $x = r - 2t$, $y = s + 2t$, $z = t$; $q(r, s, t) = 2r^2 - 3s^2 + 29t^2$;
(c) $x = r - s - t$, $y = s - t$, $z = t$; $q(r, s, t) = r^2 - 2s^2$

12.37 $q(x, y) = x^2 - y^2$, $u = (1, 1)$, $v = (1, -1)$

12.40 (a) É, (b) Não é, (c) Não é, (d) É

12.41 (a) $k > \frac{25}{8}$, (b) $-12 < k < 12$, (c) $k > 5$

12.44 (a) $P = [1, i; \ 0, 1]$, $D = I$, $s = 2$; (b) $P = [1, -2 + 3i; \ 0, 1]$, $D = \text{diag}(1, -14)$, $s = 0$;
(c) $P = [1, i, -3 + i; \ 0, 1, i; \ 0, 0, 1]$, $D = \text{diag}(1, 1, -4)$, $s = 1$

Operadores Lineares em Espaços com Produto Interno

13.1 INTRODUÇÃO

Neste capítulo investigamos o espaço $A(V)$ dos operadores lineares T de um espaço com produto interno V. (Ver Capítulo 7.) Assim, o corpo K ou é o corpo dos reais $\mathbf{R}$ ou é o corpo dos complexos $\mathbf{C}$. Na verdade, utilizamos terminologias diferentes para o caso real e o caso complexo. Também utilizamos o fato de que os produtos internos dos espaços euclidianos real $\mathbf{R}^n$ e complexo $\mathbf{C}^n$ poderem ser definidos, respectivamente, por

$$\langle u, v\rangle = u^T v \qquad \text{e} \qquad \langle u, v\rangle = u^T \bar{v}$$

onde u e v são vetores coluna.

O leitor deveria rever o conteúdo do Capítulo 7 para estar bastante familiarizado com os conceitos de norma (comprimento), ortogonalidade e bases ortonormais. Observamos, também, que no Capítulo 7 tratamos, principalmente, com espaços com produtos internos reais, ao passo que aqui vamos supor que V é um espaço com produto interno complexo, salvo menção em contrário.

Finalmente, observamos que no Capítulo 2 denotamos a transposta conjugada de uma matriz complexa A por A^H, isto é, $A^H = \overline{A^T}$. Essa notação não é padrão. Muitos textos, principalmente os mais avançados, usam A^* para denotar essa matriz. Passamos a utilizar essa notação neste capítulo. Assim, agora, $A^* = \overline{A^T}$.

13.2 OPERADORES ADJUNTOS

Começamos com uma definição básica.

DEFINIÇÃO Dizemos que um operador linear T de um espaço com produto interno V tem um *operador adjunto* T^* de V se $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, para quaisquer $u, v \in V$.

No exemplo a seguir, mostramos que o operador adjunto tem uma descrição simples no contexto das transformações matriciais.

Exemplo 13.1

(a) Seja A uma matriz real quadrada de ordem n vista como um operador linear de $\mathbf{R}^n$ Então, dados $u, v \in \mathbf{R}_n$ quaisquer,

$$\langle Au, v\rangle = (Au)^T v = u^T A^T v = \langle u, A^T v\rangle$$

Assim, a transposta A^T de A é a adjunta de A.

(b) Seja B uma matriz complexa quadrada de ordem n vista como um operador linear de $\mathbf{C}^n$. Então, dados $u, v, \in \mathbf{C}^n$ quaisquer,

$$\langle Bu, v\rangle = (Bu)^T \bar{v} = u^T B^T \bar{v} = u^T \overline{B^*} \bar{v} = \langle u, B^* v\rangle$$

Assim, a transposta conjugada B^* de B é a adjunta de B.

OBSERVAÇÃO A notação B^* pode significar tanto a adjunta de B como um operador linear ou a transposta conjugada de B como uma matriz. Pela parte (b) do exemplo precedente, essa ambiguidade é irrelevante, pois ambas denotam o mesmo objeto.

O teorema seguinte (demonstrado no Problema 13.4) é o resultado principal desta seção.

Teorema 13.1 Seja T um operador linear de um espaço vetorial com produto interno V de dimensão finita sobre K. Valem as afirmações a seguir.

(i) Existe um único operador linear T^* de V tal que $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, para quaisquer $u, v \in V$. (Ou seja, T tem um adjunto T^*.)

(ii) Se A é a representação matricial de T em relação a uma base ortonormal $S = \{u_i\}$ de V, então a representação matricial de T^* na base S é a transposta conjugada A^* de A (ou a transposta A^T de A se K for real).

Enfatizamos que essa relação simples entre as representações matriciais de T e de T^* não existe se S não for ortonormal. Assim, vemos uma propriedade útil das bases ortonormais. Também enfatizamos que esse teorema não é válido se V tiver dimensão infinita (Problema 13.31).

O teorema seguinte (demonstrado no Problema 13.5) resume algumas das propriedades do adjunto.

Teorema 13.2 Sejam T, T_1, T_2 operadores lineares de V e seja $k \in K$. Então

(i) $(T_1 + T_2)^* = T_1^* + T_2^*$, (iii) $(T_1T_2)^* = T_2^*T_1^*$,

(ii) $(kT)^* = \bar{k}T^*$, (iv) $(T^*)^* = T$.

Observe a semelhança entre esse teorema e o Teorema 2.3 relativo às propriedades da operação de transposição de matrizes.

Funcionais lineares e espaços com produto interno

No Capítulo 11 vimos que um funcional linear ϕ de um espaço vetorial V é uma transformação linear $\phi: V \to K$. Nessa subseção apresentamos um resultado básico (Teorema 13.3) que é utilizado na demonstração do importante Teorema 13.1.

Seja V um espaço com produto interno. Cada $u \in V$ determina uma aplicação $\hat{u}: V \to K$ definida por

$$\hat{u}(v) = \langle v, u\rangle$$

Agora, dados quaisquer $a, b \in K$ e $v_1, v_2 \in V$, temos

$$\hat{u}(av_1 + bv_2) = \langle av_1 + bv_2, \ u\rangle = a\langle v_1, u\rangle + b\langle v_2, u\rangle = a\hat{u}(v_1) + b\hat{u}(v_2)$$

Assim, $\hat{u}$ é um funcional linear de V. A recíproca também é verdadeira em espaços de dimensão finita e está incluída no importante teorema a seguir (demonstrado no Problema 13.3).

Teorema 13.3 Seja ϕ um funcional linear de um espaço com produto interno V de dimensão finita. Então existe um único vetor $u \in V$ tal que $\phi(v) = \langle v, u\rangle$, para cada $v \in V$.

Observamos que esse teorema não é válido em espaços de dimensão infinita (Problema 13.24).

13.3 ANALOGIA ENTRE $A(V)$ E $\mathbf{C}$, OPERADORES LINEARES ESPECIAIS

Seja $A(V)$ a álgebra de todos os operadores lineares de um espaço com produto interno V de dimensão finita. A aplicação adjunta $T \mapsto T^*$ de $A(V)$ é muito parecida com a aplicação de conjugação $z \mapsto \bar{z}$ do corpo dos complexos $\mathbf{C}$. Para ilustrar essa analogia, identificamos na Tabela 13-1 certas classes de operadores $T \in A(V)$ cujo comportamento pela aplicação adjunta imita o comportamento pela conjugação de certas classes de números complexos.

A analogia entre esses operadores T e os números complexos z é refletida no próximo teorema.

Tabela 13.1

Classe de números complexos	Comportamento pela conjugação	Classe de operadores de $A(V)$	Comportamento pela aplicação adjunta
Círculo unitário ($\|z\| = 1$)	$\bar{z} = 1/z$	Operadores ortogonais (caso real) Operadores unitários (caso complexo)	$T^* = T^{-1}$
Eixo real	$\bar{z} = z$	Operadores autoadjuntos Também denominados simétricos (caso real) ou hermitianos (caso complexo)	$T^* = T$
Eixo imaginário	$\bar{z} = -z$	Operadores antiadjuntos Também denominados antissimétricos (caso real) anti-hermitianos ou (caso complexo)	$T^* = -T$
Eixo real positivo $(0, \infty)$	$z = \bar{w}w, w \neq 0$	Operadores positivos	$T = S^*S$, com S não singular

Teorema 13.4 Seja λ um autovalor de um operador linear T de V.

(i) Se $T^* = T^{-1}$ (ou seja, T é ortogonal ou unitário), então $|\lambda| = 1$.
(ii) Se $T^* = T$ (ou seja, T é autoadjunto), então λ é real.
(iii) Se $T^* = -T$ (ou seja, T é antiadjunto), então λ é imaginário puro.
(iv) Se $T = S^*S$ com S não singular (ou seja, T é positivo), então λ é real e positivo.

Demonstração. Em cada caso, seja v um autovetor não nulo de T associado a λ; ou seja, $T(v) = \lambda v$, com $v \neq 0$. Logo, $\langle v, v\rangle$ é positivo.

Demonstração de (i). Mostramos que $\lambda\bar{\lambda}\langle v, v\rangle = \langle v, v\rangle$:

$$\lambda\bar{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, I(v)\rangle = \langle v, v\rangle$$

Como $\langle v, v\rangle \neq 0$; temos $\lambda\bar{\lambda} = 1$ e, portanto, $|\lambda| = 1$.

Demonstração de (ii). Mostramos que $\lambda\langle v, v\rangle = \bar{\lambda}\langle v, v\rangle$:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle$$

Como $\langle v, v\rangle \neq 0$; temos $\lambda = \bar{\lambda}$ e, portanto, λ é real.

Demonstração de (iii). Mostramos que $\lambda\langle v, v\rangle = -\bar{\lambda}\langle v, v\rangle$:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, -T(v)\rangle = \langle v, -\lambda v\rangle = -\bar{\lambda}\langle v, v\rangle$$

Como $\langle v, v\rangle \neq 0$; temos $\lambda = -\bar{\lambda}$ ou $\bar{\lambda} = -\lambda$ e, portanto, λ é imaginário puro.

Demonstração de (iv). Observe, inicialmente, que $S(v) \neq 0$, pois S é não singular. Logo, $\langle S(v), S(v)\rangle$ é positivo. Mostramos que $\lambda\langle v, v\rangle = \langle S(v), S(v)\rangle$.

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle S^*S(v), v\rangle = \langle S(v), S(v)\rangle$$

Como $\langle v, v\rangle$ e $\langle S(v), S(v)\rangle$ são positivos, temos que λ é positivo.

OBSERVAÇÃO Cada um desses operadores T comuta com seu adjunto, isto é, $TT^* = T^*T$. Tais operadores são denominados *operadores normais*.

13.4 OPERADORES AUTOADJUNTOS

Seja T um operador autoadjunto de um espaço com produto interno V, isto é,

$$T = T^*$$

(Se T estiver dado por uma matriz A, então A é simétrica ou hermitiana de acordo com A ser real ou complexa.) Pelo Teorema 13.4, os autovalores de T são reais. Uma outra propriedade importante de T é dada no teorema seguinte.

Teorema 13.5 Sejam T um operador autoadjunto de V e u e v autovetores de T associados a autovalores distintos. Então u e v são ortogonais, isto é, $\langle u, v\rangle = 0$.

Demonstração. Digamos que $T(u) = \lambda_1 u$ e $T(v) = \lambda_2 v$, com $\lambda_1 \neq \lambda_2$. Mostramos que $\lambda_1\langle u, v\rangle = \lambda_2\langle u, v\rangle$.

$$\begin{aligned}\lambda_1\langle u, v\rangle &= \langle \lambda_1 u, v\rangle = \langle T(u), v\rangle = \langle u, T^*(v)\rangle = \langle u, T(v)\rangle \\ &= \langle u, \lambda_2 v\rangle = \bar{\lambda}_2\langle u, v\rangle = \lambda_2\langle u, v\rangle\end{aligned}$$

(A quarta igualdade decorre de $T^* = T$ e a última usa o fato de que o autovalor λ_2 é real.) Como $\lambda_1 \neq \lambda_2$, obtemos $\langle u, v\rangle = 0$, demonstrando o teorema.

13.5 OPERADORES ORTOGONAIS E UNITÁRIOS

Seja U um operador linear de um espaço com produto interno V de dimensão finita tal que

$$U^* = U^{-1} \qquad \text{ou, equivalentemente,} \qquad UU^* = U^*U = I$$

Já vimos que U é denominado ortogonal ou unitário, dependendo de o corpo subjacente ser real ou complexo. O teorema seguinte (demonstrado no Problema 13.10) dá uma caracterização alternativa desses operadores.

Teorema 13.6 As condições seguintes sobre um operador U são equivalentes.

(i) $U^* = U^{-1}$; isto é, $UU^* = U^*U = I$. [U é unitário (ortogonal).]

(ii) U preserva produtos internos, isto é, dados $v, w \in V$, quaisquer, $\langle U(v),\ U(w)\rangle = \langle v, w\rangle$.

(iii) U preserva comprimentos, isto é, $\|U(v)\| = \|v\|$, para cada $v \in V$.

Exemplo 13.2

(a) Seja $T: \mathbf{R}^3 \to \mathbf{R}^3$ o operador linear que gira cada vetor v horizontalmente em torno do eixo z por um ângulo θ fixado, conforme Figura 10-1 (Seção 10.3). Esse operador é definido por

$$T(x, y, z) = (x\cos\theta - y\,\text{sen}\,\theta,\ \ x\,\text{sen}\,\theta + y\cos\theta,\ \ z)$$

Observamos que os comprimentos (distâncias a partir da origem) são preservados por T. Assim, T é um operador ortogonal.

(b) Seja V o espaço de Hilbert l_2 definido na Seção 7.3. Seja $T: V \to V$ o operador linear definido por

$$T(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots)$$

Claramente, T preserva produtos internos e comprimentos. Contudo, T não é uma aplicação sobrejetora porque, por exemplo, (1, 0, 0, ...) não pertence à imagem de T; logo, T não é invertível. Vemos, assim, que o Teorema 13.6 não é válido em espaços de dimensão infinita.

Um *isomorfismo* de um espaço com produto interno sobre um outro é uma aplicação bijetora que preserva as três propriedades básicas de um espaço com produto interno: a soma de vetores, a multiplicação por escalar e o produto interno. Assim, as transformações ortogonais e unitárias também podem ser caracterizadas como os isomorfismos de V sobre si mesmo. Observe que uma transformação U dessas também preserva distâncias, já que

$$\|U(v) - U(w)\| = \|U(v - w)\| = \|v - w\|$$

Por isso, também dizemos que U é uma *isometria*.

13.6 MATRIZES ORTOGONAIS E UNITÁRIAS

Seja U um operador linear de um espaço com produto interno V. A partir do Teorema 13.1, obtemos os resultados a seguir.

Teorema 13.7a Uma matriz complexa A representa um operador unitário U (em relação a alguma base ortonormal) se, e só se, $A^* = A^{-1}$.

Teorema 13.7b Uma matriz real A representa um operador ortogonal U (em relação a alguma base ortonormal) se, e só se, $A^T = A^{-1}$.

Esses teoremas motivam as definições seguintes (apresentadas nas Seções 2.10 e 2.11).

DEFINIÇÃO Uma matriz complexa A tal que $A^* = A^{-1}$ é denominada *matriz unitária*.

DEFINIÇÃO Uma matriz real A tal que $A^T = A^{-1}$ é denominada *matriz ortogonal*.

Repetimos o Teorema 2.6, que caracteriza essas matrizes.

Teorema 13.8 Seja A uma matriz. As afirmações seguintes são equivalentes.

(i) A é unitária (ortogonal).
(ii) As linhas de A constituem um conjunto ortonormal.
(iii) As colunas de A constituem um conjunto ortonormal.

13.7 MUDANÇA DE BASES ORTONORMAIS

As bases ortonormais desempenham um papel especial na teoria dos espaços vetoriais com produto interno V. Assim, naturalmente estamos interessados nas propriedades da matriz de mudança de base de uma dessas bases para outra. Vale o teorema seguinte (demonstrado no Problema 13.12).

Teorema 13.9 Seja $\{u_1, ..., u_n\}$ uma base ortonormal de um espaço com produto interno V. Então, a matriz de mudança de base de $\{u_i\}$ para uma outra base ortonormal é unitária (ortogonal). Reciprocamente, se $P = [a_{ij}]$ é uma matriz unitária (ortogonal), então o conjunto dado a seguir é uma base ortonormal de V.

$$\{u_i' = a_{1i}u_1 + a_{2i}u_2 + \cdots + a_{ni}u_n : i = 1, \ldots, n\}$$

Sabemos que matrizes A e B que representam um mesmo operador linear T são semelhantes, isto é, $B = P^{-1}AP$, onde P é a matriz (não singular) de mudança de base. Por outro lado, se V é um espaço com produto interno, estamos usualmente interessados no caso em que P é unitária (ou ortogonal), como foi sugerido no Teorema 13.9. (Lembremos que P é unitária se sua inversa for a transposta conjugada $P^* = P^{-1}$ e que P é ortogonal se sua inversa for a transposta $P^T = P^{-1}$.) Isso conduz à seguinte definição.

DEFINIÇÃO As matrizes complexas A e B são *unitariamente equivalentes* se existir alguma matriz unitária P tal que $B = P^*AP$. Analogamente, as matrizes reais A e B são *ortogonalmente equivalentes* se existir alguma matriz ortogonal P tal que $B = P^TAP$.

Observe que matrizes ortogonalmente equivalentes são necessariamente congruentes.

13.8 OPERADORES NÃO NEGATIVOS E POSITIVOS

Seja P um operador linear de um espaço com produto interno V. Dizemos que

(i) P é *positivo* se $P = S^*S$, para algum operador não singular S.
(ii) P é *não negativo* se $P = S^*S$, para algum operador S.

Uma caracterização alternativa desses operadores é dada nos teoremas seguintes.

Teorema 13.10a Seja P um operador. As afirmações seguintes são equivalentes.

(i) $P = T^2$, para algum operador não singular autoadjunto T.
(ii) P é positivo.
(iii) P é autoadjunto e $\langle P(u), u \rangle > 0$, para cada $u \neq 0$ de V.

Vale o teorema correspondente para operadores não negativos (demonstrado no Problema 13.21).

Teorema 13.10b Seja P um operador. As afirmações seguintes são equivalentes.

(i) $P = T^2$, para algum operador autoadjunto T.
(ii) P é não negativo, isto é, $P = S^*S$, para algum operador S.
(iii) P é autoadjunto e $\langle P(u), u \rangle \geq 0$, para cada $u \in V$ de V.

13.9 DIAGONALIZAÇÃO E FORMAS CANÔNICAS EM ESPAÇOS COM PRODUTO INTERNO

Seja T um operador linear de um espaço com produto interno V de dimensão finita sobre o corpo K. A representação de T por uma matriz diagonal depende dos autovetores e autovalores de T e, portanto, depende das raízes do polinômio característico $\Delta(t)$ de T. Agora, $\Delta(t)$ sempre se fatora em polinômios lineares sobre o corpo complexo $\mathbf{C}$, mas pode não conter fatores lineares sobre o corpo real $\mathbf{R}$. Assim, a situação para espaços com produto interno reais (também denominados espaços euclidianos) é intrinsecamente diferente da situação para espaços com produto interno complexos (também denominados espaços unitários) e, portanto, serão tratados separadamente.

Espaços com produto interno reais, operadores simétricos e ortogonais

Vale o teorema seguinte (demonstrado no Problema 13.14).

Teorema 13.11 Seja T um operador simétrico (autoadjunto) de um espaço com produto interno real V de dimensão finita. Então existe uma base ortonormal de V constituída por autovetores de T, isto é, T pode ser representado por uma matriz diagonal em relação a alguma base ortonormal.

Segue a afirmação correspondente para matrizes.

Teorema 13.11 (Forma Alternativa) Seja A uma matriz simétrica real. Então existe uma matriz ortogonal P tal que $B = P^{-1}AP = P^TAP$ é diagonal.

Podemos escolher as colunas da matriz P do enunciado como autovetores ortogonais normalizados de A; então, os elementos diagonais de B são os autovalores associados.

Por outro lado, um operador ortogonal T não precisa ser simétrico, logo, pode não ser representável por uma matriz diagonal em relação a uma base ortonormal. Entretanto, um operador T desses possui uma representação canônica simples, conforme descrito no teorema a seguir (demonstrado no Problema 13.16).

Teorema 13.12 Seja T um operador ortogonal de um espaço com produto interno real V. Então existe uma base ortonormal de V na qual T é representado por uma matriz diagonal em blocos M do tipo

$$M = \operatorname{diag}\left(I_s, \quad -I_t, \quad \begin{bmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 \\ \operatorname{sen}\theta_1 & \cos\theta_1 \end{bmatrix}, \quad \ldots, \quad \begin{bmatrix} \cos\theta_r & -\operatorname{sen}\theta_r \\ \operatorname{sen}\theta_r & \cos\theta_r \end{bmatrix}\right)$$

Observemos que cada um dos blocos diagonais 2×2 representa uma rotação num subespaço de dimensão dois associado e que cada elemento diagonal –1 da matriz $-I_t$ oposta da identidade representa uma reflexão num subespaço unidimensional associado.

Espaços com produto interno complexos, operadores normais e triangulares

Dizemos que um operador linear T é *normal* se comutar com seu adjunto, isto é, se $TT^* = T^*T$. Observe que os operadores normais incluem ambos os operadores autoadjuntos e os unitários.

Analogamente, dizemos que uma matriz complexa A é *normal* se comutar com sua transposta conjugada, isto é, se $AA^* = A^*A$.

Exemplo 13.3 Seja $A = \begin{bmatrix} 1 & 1 \\ i & 3+2i \end{bmatrix}$. Então $A^* = \begin{bmatrix} 1 & -i \\ 1 & 3-2i \end{bmatrix}$ e, também, $AA^* = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 14 \end{bmatrix} = A^*A$. Assim, A é normal.

Vale o teorema seguinte (demonstrado no Problema 13.19).

Teorema 13.13 Seja T um operador normal de um espaço com produto interno complexo V de dimensão finita. Então existe uma base ortonormal de V constituída por autovetores de T, isto é, T pode ser representado por uma matriz diagonal em relação a alguma base ortonormal.

Segue a afirmação correspondente para matrizes.

Teorema 13.13 (Forma Alternativa) Seja A uma matriz normal. Então existe uma matriz unitária P tal que $B = P^{-1}AP = P^*AP$ é diagonal.

No teorema a seguir (demonstrado no Problema 13.20) vemos que mesmo operadores de espaços unitários que não sejam normais têm uma forma relativamente simples.

Teorema 13.14 Seja T um operador qualquer de um espaço com produto interno complexo V de dimensão finita. Então T pode ser representado por uma matriz triangular em relação a alguma base ortonormal de V.

Teorema 13.14 (Forma Alternativa) Seja A uma matriz complexa qualquer. Então existe uma matriz unitária P tal que $B = P^{-1}AP = P^*AP$ é triangular.

13.10 TEOREMA ESPECTRAL

O teorema espectral é uma reformulação dos teoremas de diagonalização 13.11 e 13.13.

Teorema 13.15 (Teorema Espectral) Seja T um operador normal (simétrico) de um espaço com produto interno V complexo (real) de dimensão finita. Então existem operadores lineares $E_1, \ldots, E_r$ de V e escalares $\lambda_1, \ldots, \lambda_r$ tais que

(i) $T = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_r E_r$, (iii) $E_1^2 = E_1, E_2^2 = E_2, \ldots, E_r^2 = E_r$,

(ii) $E_1 + E_2 + \cdots + E_r = I$, (iv) $E_i E_j = 0$ para $i \neq j$.

Esses operadores lineares E_1, ..., E_r de V são *projeções*, isto é, $E_i^2 = E_i$. Além disso, são *projeções ortogonais*, porque têm a propriedade adicional que $E_iE_j = 0$, com $i \neq j$.

No exemplo a seguir, vemos a relação entre a representação matricial diagonal e as correspondentes projeções ortogonais.

Exemplo 13.4 Considere as matrizes diagonais A, E_1, E_2, E_3 seguintes.

$$A = \begin{bmatrix} 2 & & & \\ & 3 & & \\ & & 3 & \\ & & & 5 \end{bmatrix}, \quad E_1 = \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{bmatrix}$$

Temos

(i) $A = 2E_1 + 3E_2 + 5E_3$, (ii) $E_1 + E_2 + E_3 = I$, (iii) $E_i^2 = E_i$, (iv) $E_iE_j = 0$, com $i \neq j$.

Problemas Resolvidos

Operadores adjuntos

13.1 Encontre o adjunto do operador $F:\mathbf{R}^3 \to \mathbf{R}^3$ definido por

$$F(x, y, z) = (3x + 4y - 5z, \quad 2x - 6y + 7z, \quad 5x - 9y + z)$$

Começamos calculando a matriz A que representa F na base canônica de $\mathbf{R}^3$, ou seja, a matriz A cujas colunas são os coeficientes de x, y, z e, depois, tomamos a transposta A^T de A, como segue.

$$A = \begin{bmatrix} 3 & 4 & -5 \\ 2 & -6 & 7 \\ 5 & -9 & 1 \end{bmatrix} \quad \text{e, então,} \quad A^T = \begin{bmatrix} 3 & 2 & 5 \\ 4 & -6 & -9 \\ -5 & 7 & 1 \end{bmatrix}$$

A adjunta F^* é representada pela transposta de A; logo,

$$F^*(x, y, z) = (3x + 2y + 5z, \quad 4x - 6y - 9z, \quad -5x + 7y + z)$$

13.2 Encontre o adjunto do operador $G:\mathbf{C}^3 \to \mathbf{C}^3$ definido por

$$G(x, y, z) = [2x + (1 - i)y, \quad (3 + 2i)x - 4iz, \quad 2ix + (4 - 3i)y - 3z]$$

Começamos calculando a matriz B que representa G na base canônica de $\mathbf{C}^3$ e, depois, tomamos a transposta conjugada B^* de B, como segue.

$$B = \begin{bmatrix} 2 & 1 - i & 0 \\ 3 + 2i & 0 & -4i \\ 2i & 4 - 3i & -3 \end{bmatrix} \quad \text{e, então,} \quad B^* = \begin{bmatrix} 2 & 3 - 2i & -2i \\ 1 + i & 0 & 4 + 3i \\ 0 & 4i & -3 \end{bmatrix}$$

Assim, $G^*(x, y, z) = [2x + (3 - 2i)y - 2iz, \quad (1 + i)x + (4 + 3i)z, \quad 4iy - 3z]$.

13.3 Demonstre o Teorema 13.3. Seja ϕ um funcional linear de um espaço com produto interno V de dimensão finita. Então existe um único vetor $u \in V$ tal que $\phi(v) = \langle v, u \rangle$, para cada $v \in V$.

Seja $\{w_1, ..., w_n\}$ uma base ortonormal de V. Definimos

$$u = \overline{\phi(w_1)}w_1 + \overline{\phi(w_2)}w_2 + \cdots + \overline{\phi(w_n)}w_n$$

Seja $\hat{u}$ o funcional linear de V definido por $\hat{u}(v) = \langle v, u \rangle$, para cada $v \in V$. Então, para $i = 1, ..., n$,

$$\hat{u}(w_i) = \langle w_i, u \rangle = \langle w_i, \quad \overline{\phi(w_1)}w_1 + \cdots + \overline{\phi(w_n)}w_n \rangle = \phi(w_i)$$

Como $\hat{u}$ e ϕ coincidem nos vetores da base, obtemos $\hat{u} = \phi$.
Suponha, agora, que u' seja um outro vetor de V tal que $\phi(v) = \langle v, u' \rangle$, para cada $v \in V$. Então $\langle v, u \rangle = \langle v, u' \rangle$, ou $\langle v, \ u - u' \rangle = 0$. Em particular, isso vale para o vetor $v = u - u'$, ou seja, $\langle u - u', \ u - u' \rangle = 0$. Disso decorre que $u - u' = 0$ e $u = u'$. Assim, o vetor u é único, como queríamos mostrar.

13.4 Demonstre o Teorema 13.1. Seja T um operador linear de um espaço vetorial com produto interno V de dimensão finita. Valem as afirmações a seguir.

(a) Existe um único operador linear T^* de V tal que

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle, \text{ para quaisquer } u, v \in V.$$

(b) Se A é a representação matricial de T em relação a uma base ortonormal $S = \{u_i\}$ de V, então a representação matricial de T^* na base S é a transposta conjugada A^* de A.

(a) Definamos o operador T^*. Seja v um vetor arbitrário de V fixado. A aplicação $u \mapsto \langle T(u), v \rangle$ é um funcional linear de V. Logo, pelo Teorema 13.3, existe um único vetor $v' \in V$ tal que $\langle T(u), v \rangle = \langle u, v' \rangle$, para cada $u \in V$. Definimos $T^* : V \to V$ por $T^*(v) = v'$. Então $\langle T(u), v \rangle = \langle u, T^*(v) \rangle$, para cada $u, v \in V$.
Mostremos que T^* é linear. Dados $u, v_i \in V$, e $a, b \in K$, quaisquer,

$$\begin{aligned}\langle u, \ T^*(av_1 + bv_2) \rangle &= \langle T(u), \ av_1 + bv_2 \rangle = \bar{a}\langle T(u), v_1 \rangle + \bar{b}\langle T(u), v_2 \rangle \\ &= \bar{a}\langle u, T^*(v_1) \rangle + \bar{b}\langle u, T^*(v_2) \rangle = \langle u, aT^*(v_1) + bT^*(v_2) \rangle\end{aligned}$$

Como isso é válido para cada $u \in V$, temos $T^*(av_1 + bv_2) = aT^*(v_1) + bT^*(v_2)$. Assim, T^* é linear.

(b) As matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$ que representam T e T^*, respectivamente, na base S são dadas por $a_{ij} = \langle T(u_j), u_i \rangle$ e $b_{ij} = \langle T^*(u_j), u_i \rangle$ (Problema 13.67). Logo,

$$b_{ij} = \langle T^*(u_j), u_i \rangle = \overline{\langle u_i, T^*(u_j) \rangle} = \overline{\langle T(u_i), u_j \rangle} = \overline{a_{ji}}$$

Assim, $B = A^*$, como queríamos mostrar.

13.5 Demonstre o Teorema 13.2.

(i) $(T_1 + T_2)^* = T_1^* + T_2^*$, (iii) $(T_1T_2)^* = T_2^*T_1^*$,
(ii) $(kT)^* = \bar{k}T^*$, (iv) $(T^*)^* = T$.

(i) Dados $u, v \in V$,

$$\begin{aligned}\langle (T_1 + T_2)(u), v \rangle &= \langle T_1(u) + T_2(u), \ v \rangle = \langle T_1(u), v \rangle + \langle T_2(u), v \rangle \\ &= \langle u, T_1^*(v) \rangle + \langle u, T_2^*(v) \rangle = \langle u, \ T_1^*(v) + T_2^*(v) \rangle \\ &= \langle u, \ (T_1^* + T_2^*)(v) \rangle\end{aligned}$$

Pela unicidade do adjunto, resulta $(T_1 + T_2)^* = T_1^* + T_2^*$.

(ii) Dados $u, v \in V$,

$$\langle (kT)(u), \ v \rangle = \langle kT(u), v \rangle = k\langle T(u), v \rangle = k\langle u, T^*(v) \rangle = \langle u, \bar{k}T^*(v) \rangle = \langle u, (\bar{k}T^*)(v) \rangle$$

A unicidade do adjunto implica $(kT)^* = \bar{k}T^*$.

(iii) Dados $u, v \in V$,

$$\begin{aligned}\langle (T_1T_2)(u), v \rangle &= \langle T_1(T_2(u)), v \rangle = \langle T_2(u), T_1^*(v) \rangle \\ &= \langle u, T_2^*(T_1^*(v)) \rangle = \langle u, (T_2^*T_1^*)(v) \rangle\end{aligned}$$

A unicidade do adjunto implica $(T_1T_2)^* = T_2^*T_1^*$.

(iv) Dados $u, v \in V$,

$$\langle T^*(u), v \rangle = \overline{\langle v, T^*(u) \rangle} = \overline{\langle T(v), u \rangle} = \langle u, T(v) \rangle$$

A unicidade do adjunto implica $(T^*)^* = T$.

13.6 Mostre que (a) $I^* = I$, (b) $0^* = 0$.

(a) Dados $u, v \in V$, $\langle I(u), v\rangle = \langle u, v\rangle = \langle u, I(v)\rangle$; logo, $I^* = I$.

(b) Dados $u, v \in V$, $\langle 0(u), v\rangle = \langle 0, v\rangle = 0 = \langle u, 0\rangle = \langle u, 0(v)\rangle$; logo, $0^* = 0$.

13.7 Seja T invertível. Mostre que $(T^{-1})^* = (T^*)^{-1}$.

$$I = I^* = (TT^{-1})^* = (T^{-1})^*T^*; \text{ portanto, } (T^{-1})^* = (T^*)^{-1}.$$

13.8 Sejam T um operador linear de V e W um subespaço T-invariante de V. Mostre que $W^\perp$ é invariante por T^*.

Seja $u \in W^\perp$. Se $w \in W$, então $T(w) \in W$ e, portanto, $\langle w, T^*(u)\rangle = \langle T(w), u\rangle = 0$. Logo, $T^*(u) \in W^\perp$, já que é ortogonal a cada $w \in W$. Assim, $W^\perp$ é invariante por T^*.

13.9 Seja T um operador linear de V. Mostre que cada uma das condições seguintes implica $T = 0$.

(i) $\langle T(u), v\rangle = 0$, para quaisquer $u, v \in V$.

(ii) V é um espaço complexo e $\langle T(u), u\rangle = 0$, para cada $u \in V$.

(iii) V é autoadjunto e $\langle T(u), u\rangle = 0$, para cada $u \in V$.

Dê um exemplo de um operador T de um espaço real V tal que $\langle T(u), u\rangle = 0$ valha para cada $u \in V$, mas $T \neq 0$. [Ou seja, (ii) não vale em espaços reais V.]

(i) Tomando $v = T(u)$, temos $\langle T(u), T(u)\rangle = 0$ e, portanto, $T(u) = 0$, para cada $u \in V$. Assim, $T = 0$.

(ii) Por hipótese, $\langle T(v+w), \ v+w\rangle = 0$ para quaisquer $v, w \in V$. Expandindo e usando $\langle T(v), v\rangle = 0$ e $\langle T(w), w\rangle = 0$, obtemos

$$\langle T(v), w\rangle + \langle T(w), v\rangle = 0 \tag{1}$$

Observe que w é arbitrário em (1). Substituindo w por iw e usando $\langle T(v), iw\rangle = \bar{i}\langle T(v), w\rangle = -i\langle T(v), w\rangle$ e $\langle T(iw), v\rangle = \langle iT(w), v\rangle = i\langle T(w), v\rangle$, obtemos

$$-i\langle T(v), w\rangle + i\langle T(w), v\rangle = 0$$

Dividindo por i e somando com (1), resulta $\langle T(w), v\rangle = 0$ para quaisquer $v, w, \in V$. Por (i), $T = 0$.

(iii) Por (ii), a afirmação vale em espaços complexos, portanto, resta provar a afirmação no caso real. Expandindo $\langle T(v+w), \ v+w\rangle = 0$, novamente obtemos (1). Como T é autoadjunto e o espaço é real, temos $\langle T(w), v\rangle = \langle w, T(v)\rangle = \langle T(v), w\rangle$. Substituindo isso em (1), resulta $\langle T(v), w\rangle = 0$, para quaisquer $v, w \in V$. Por (i), $T = 0$.

Como um exemplo, considere o operador linear T de $\mathbf{R}^2$ definido por $T(x, y) = (y, -x)$. Então, $\langle T(u), u\rangle = 0$, para cada $u \in V$, mas $T \neq 0$.

Operadores e matrizes ortogonais e unitários

13.10 Demonstre o Teorema 13.6. As condições seguintes sobre um operador U são equivalentes.

(i) $U^* = U^{-1}$, isto é, U é unitário. (ii) $\langle U(v), U(w)\rangle = \langle u, w\rangle$. (iii) $\|U(v)\| = \|v\|$

Suponha que (i) valha. Então, dados quaisquer $v, w, \in V$,

$$\langle U(v), U(w)\rangle = \langle v, U^*U(w)\rangle = \langle v, I(w)\rangle = \langle v, w\rangle$$

Assim, (i) implica (ii). Suponha que (ii) valha. Então

$$\|U(v)\| = \sqrt{\langle U(v), U(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$$

Assim, (ii) implica (iii). Resta mostrar que (iii) implica (i).

Suponha que (iii) valha. Então, para cada $v \in V$,

$$\langle U^*U(v)\rangle = \langle U(v), U(v)\rangle = \langle v, v\rangle = \langle I(v), v\rangle$$

Logo, $\langle (U^*U - I)(v), \ v\rangle = 0$, para cada $v \in V$. Como $U^*U - I$ é autoadjunto (demonstre), pelo Problema 13.9 temos que $U^*U - I = 0$ e, portanto, $U^*U = I$. Assim, $U^* = U^{-1}$, como queríamos provar.

13.11 Sejam U um operador unitário (ortogonal) de V e W um subespaço de V invariante por U. Mostre que $W^\perp$ também é invariante por U.

Como U é não singular, $U(W) = W$, de modo que, para cada $w \in W$, existe algum $w' \in W$ tal que $U(w') = w$. Seja, agora, $v \in W^\perp$. Então, para cada $w \in W$,

$$\langle U(v), w\rangle = \langle U(v), U(w')\rangle = \langle v, w'\rangle = 0$$

Assim, $U(v)$ pertence a $W^\perp$. Dessa forma, $W^\perp$ é invariante por U.

13.12 Demonstre o Teorema 13.9. A matriz de mudança de base de uma base ortonormal $\{u_1, ..., u_n\}$ para uma outra base ortonormal é unitária (ortogonal). Reciprocamente, se $P = [a_{ij}]$ é uma matriz unitária (ortogonal), então os vetores $u_{i'} = \sum_j a_{ji}u_j$ formam uma base ortonormal.

Seja $\{v_i\}$ uma outra base ortonormal e suponha que

$$v_i = b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n, \quad i = 1, \ldots, n \tag{1}$$

Como $\{v_i\}$ é ortonormal,

$$\delta_{ij} = \langle v_i, v_j\rangle = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}} \tag{2}$$

Seja $B = [b_{ij}]$ a matriz de coeficientes em (1). (Então B^T é a matriz de mudança de base de $\{u_i\}$ para $\{v_i\}$.) Então $BB^* = [c_{ij}]$, com $c_{ij} = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}}$. Por (2), temos $c_{ij} = \delta_{ij}$ e, portanto, $BB^* = I$. Dessa forma, B e, portanto, B^* é unitária.

Falta provar que $\{u_i'\}$ é ortonormal. Pelo Problema 13.67,

$$\langle u_i', u_j'\rangle = a_{1i}\overline{a_{1j}} + a_{2i}\overline{a_{2j}} + \cdots + a_{ni}\overline{a_{nj}} = \langle C_i, C_j\rangle$$

onde C_i denota a i-ésima coluna da matriz unitária (ortogonal) $P = [a_{ij}]$. Como P é unitária (ortogonal), suas colunas são ortonormais, portanto, $\langle u_i', u_j'\rangle = \langle C_i, C_j\rangle = \delta_{ij}$. Assim, $\{u_i'\}$ é uma base ortonormal.

Operadores simétricos e formas canônicas em espaços euclidianos

13.13 Seja T um operador simétrico. Mostre que (a) o polinômio característico $\Delta(t)$ de T é um produto de polinômios lineares (sobre $\mathbf{R}$) e que (b) T tem algum autovetor não nulo.

(a) Seja A a matriz que representa T nalguma base ortonormal de V; então $A = A^T$. Seja $\Delta(t)$ o polinômio característico de A. Vendo A como um operador autoadjunto complexo, o Teorema 13.4 garante que todos autovalores de A são reais. Assim,

$$\Delta(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$$

em que todos os λ_i são reais. Em outras palavras, $\Delta(t)$ é um produto de polinômios lineares sobre $\mathbf{R}$.

(b) Por (a), T tem pelo menos um autovalor (real). Assim, T tem algum autovetor não nulo.

13.14 Demonstre o Teorema 13.11. Seja T um operador simétrico de um espaço com produto interno real V de dimensão finita. Então existe uma base ortonormal de V constituída por autovetores de T. (Logo, T pode ser representado por uma matriz diagonal em relação a alguma base ortonormal.)

A demonstração é por indução na dimensão n de V. Se $\dim V = 1$, a afirmação do teorema é válida trivialmente. Suponha, então, que $\dim V = n > 1$ e que o teorema seja válido para operadores simétricos de espaços de dimensão $n - 1$. Pelo Problema 13.13, existe algum autovetor não nulo v_1 de T. Seja W o espaço gerado por v_1 e seja u_1 um vetor unitário de W, por exemplo, $u_1 = v_1/\|v_1\|$.
Como v_1 é um autovetor de T, o subespaço W de V é invariante por T. Pelo Problema 13.8, $W^\perp$ é invariante por $T^* = T$. Assim, a restrição $\hat{T}$ de T a $W^\perp$ é um operador simétrico. Pelo Teorema 7.4, temos $V = W \oplus W^\perp$. Logo, $W^\perp = n - 1$, já que $\dim W = 1$. Por indução, existe uma base ortonormal $\{u_2, ..., u_n\}$ de $W^\perp$ consistindo de autovetores de $\hat{T}$ e, portanto, de T. Entretanto, temos $\langle u_1, u_i\rangle = 0$, para $i = 2, ..., n$, já que $u_i \in W^\perp$. Dessa forma, $\{u_1, u_2, ..., u_n\}$ é um conjunto ortonormal consistindo em autovetores de T. O teorema está demonstrado.

13.15 Seja $q(x, y) = 3x^2 - 6xy + 11y^2$. Encontre uma mudança de coordenadas ortogonal (substituição linear) que diagonalize a forma quadrática q.

Calculamos a matriz simétrica A que representa q e seu polinômio característico $\Delta(t)$, como segue.

$$A = \begin{bmatrix} 3 & -3 \\ -3 & 11 \end{bmatrix} \quad \text{e} \quad \Delta(t) = t^2 - \operatorname{tr}(A)\, t + |A| = t^2 - 14t + 24 = (t-2)(t-12)$$

Os autovalores são $\lambda = 2$ e $\lambda = 12$. Logo, a forma diagonal de q é

$$q(s,t) = 2s^2 + 12t^2$$

(em que usamos novas variáveis s e t). A mudança de coordenadas ortogonal correspondente é obtida por meio de um conjunto ortogonal de autovalores de A.

Subtraímos $\lambda = 2$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} \quad \text{correspondente a} \quad \begin{aligned} x - 3y &= 0 \\ -3x + 9y &= 0 \end{aligned} \quad \text{ou} \quad x - 3y = 0$$

Uma solução não nula é $u_1 = (3, 1)$. Em seguida, subtraímos $\lambda = 12$ das entradas diagonais de A para obter a matriz

$$M = \begin{bmatrix} -9 & -3 \\ -3 & -1 \end{bmatrix} \quad \text{correspondente a} \quad \begin{aligned} -9x - 3y &= 0 \\ -3x - \ \ y &= 0 \end{aligned} \quad \text{ou} \quad -3x - y = 0$$

Uma solução não nula é $u_2 = (-1, 3)$. Normalizamos u_1 e u_2 para obter a base ortonormal

$$\hat{u}_1 = (3/\sqrt{10},\ 1/\sqrt{10}), \qquad \hat{u}_2 = (-1/\sqrt{10},\ 3/\sqrt{10})$$

Seja P a matriz cujas colunas são $\hat{u}_1$ e $\hat{u}_2$. Então

$$P = \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix} \quad \text{e} \quad D = P^{-1}AP = P^TAP = \begin{bmatrix} 2 & 0 \\ 0 & 12 \end{bmatrix}$$

Assim, a mudança de coordenadas ortogonal é

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} s \\ t \end{bmatrix} \quad \text{ou} \quad x = \frac{3s - t}{\sqrt{10}}, \qquad y = \frac{s + 3t}{\sqrt{10}}$$

Também podemos escrever s e t em termos de x e y usando $P^{-1} = P^T$, como segue.

$$s = \frac{3x + y}{\sqrt{10}}, \qquad t = \frac{-x + 3y}{\sqrt{10}}$$

13.16 Demonstre o Teorema 13.12. Seja T um operador ortogonal de um espaço com produto interno real V. Então existe uma base ortonormal de V na qual T é representado por uma matriz diagonal em blocos M do tipo

$$M = \operatorname{diag}\left(1, \ldots, 1, -1, \ldots, -1, \begin{bmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 \\ \operatorname{sen}\theta_1 & \cos\theta_1 \end{bmatrix}, \ \ldots, \ \begin{bmatrix} \cos\theta_r & -\operatorname{sen}\theta_r \\ \operatorname{sen}\theta_r & \cos\theta_r \end{bmatrix}\right)$$

Seja $S = T + T^{-1} = T + T^*$. Então $S^* = (T + T^*)^* = T^* + T = S$. Assim, S é um operador simétrico de V. Pelo Teorema 13.11, existe uma base ortonormal de V consistindo em autovetores de S. Se $\lambda_1, \ldots, \lambda_m$ denota os autovalores distintos de S, então V pode ser decomposto na soma direta $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$, em que V_i consiste nos autovetores associados a λ_i. Afirmamos que cada V_i é invariante por T. De fato, dado $v \in V$, temos $S(v) = \lambda_i v$ e, portanto,

$$S(T(v)) = (T + T^{-1})T(v) = T(T + T^{-1})(v) = TS(v) = T(\lambda_i v) = \lambda_i T(v)$$

Isso mostra que $T(v) \in V_i$ e, portanto, V_i é invariante por T. Como os V_i são ortogonais uns aos outros, podemos restringir nossa investigação ao modo pelo qual T atua em cada V_i individualmente.

Num dado V_i, temos $(T + T^{-1})v = S(v) = \lambda_i v$. Multiplicando por T, obtemos

$$(T^2 - \lambda_i T + I)(v) = 0 \tag{1}$$

Consideramos os casos $\lambda_i = \pm 2$ e $\lambda_i \neq \pm 2$ separadamente. Se $\lambda_i = \pm 2$, então $(T \pm I)^2(v) = 0$, o que conduz a $(T \pm I)(v) = 0$, ou $T(v) = \pm v$. Assim, T restrito a esse V_i é I ou $-I$.
Se $\lambda_i \neq \pm 2$, então T não tem autovetores em V_i, pois, pelo Teorema 13.4, os únicos autovalores de T são 1 e –1. Dessa forma, para $v \neq 0$, os vetores v e $T(v)$ são linearmente independentes. Seja W o subespaço gerado por v e $T(v)$. Então W é invariante por T, pois, usando (1), obtemos

$$T(T(v)) = T^2(v) = \lambda_i T(v) - v \in W$$

Pelo Teorema 7.4, $V_i = W \oplus W^\perp$. Além disso, pelo Problema 13.8, $W^\perp$ também é invariante por T. Assim, podemos decompor V_i na soma direta de subespaços bidimensionais W_j, onde os W_j são ortogonais uns aos outros e cada W_j é invariante por T. Assim, agora podemos restringir nossa investigação ao modo pelo qual T age em cada W_j individualmente. Como $T^2 - \lambda_i T + I = 0$, o polinômio característico $\Delta(t)$ de T agindo em W_j é $\Delta(t) = t^2 - \lambda_i t + 1$. Assim, o determinante de T é 1, que é o termo constante de $\Delta(t)$. Pelo Teorema 2.7, a matriz A que representa T agindo em W_j em relação a alguma base ortonormal de W_j deve ser da forma

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

A união das bases dos vários W_j nos dá uma base ortonormal de V_i e a união das bases dos V_i dá uma base ortonormal de V, na qual a matriz que representa T é da forma procurada.

Operadores normais e formas canônicas em espaços unitários

13.17 Decida se a matriz dada é normal ou não.

(a) $A = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & i \\ 1 & 2+i \end{bmatrix}$

(a) $AA^* = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}, \qquad A^*A = \begin{bmatrix} 1 & 0 \\ -i & 1 \end{bmatrix}\begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$

Como $AA^* \neq A^*A$, a matriz A não é normal.

(b) $BB^*\begin{bmatrix} 1 & i \\ 1 & 2+i \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -i & 2-i \end{bmatrix} = \begin{bmatrix} 2 & 2+2i \\ 2-2i & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -i & 2-i \end{bmatrix}\begin{bmatrix} 1 & i \\ 1 & 2+i \end{bmatrix} = B^*B$

Como $BB^* = B^*B$, a matriz B é normal.

13.18 Seja T um operador normal. Prove as afirmações dadas.

(a) $T(v) = 0$ se, e só se, $T^*(v) = 0$. (b) $T - \lambda I$ é normal.

(c) Se $T(v) = \lambda v$, então $T^*(v) = \bar{\lambda} v$; logo, cada autovetor de T é, também, um autovetor de T^*.

(d) Se $T(v) = \lambda_1 v$ e $T(w) = \lambda_2 w$, com $\lambda_1 \neq \lambda_2$, então $\langle v, w\rangle = 0$; isto é, autovetores de T associados a autovalores distintos são ortogonais.

(a) Mostramos que $\langle T(v), T(v)\rangle = \langle T^*(v), T^*(v)\rangle$: como segue.

$$\langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, TT^*(v)\rangle = \langle T^*(v), T^*(v)\rangle$$

Logo, por $[I_3]$ da definição de produto interno na Seção 7.2, vale (a).

(b) Mostramos que $T - \lambda I$ comuta com seu adjunto.

$$\begin{aligned}(T-\lambda I)(T-\lambda I)^* &= (T-\lambda I)(T^*-\bar{\lambda} I) = TT^* - \lambda T^* - \bar{\lambda} T + \lambda\bar{\lambda} I \\ &= T^*T - \bar{\lambda} T - \lambda T^* + \bar{\lambda}\lambda I = (T^* - \bar{\lambda} I)(T - \lambda I) \\ &= (T-\lambda I)^*(T-\lambda I)\end{aligned}$$

Assim, $T - \lambda I$ é normal.

(c) Se $T(v) = \lambda v$, então $(T - \lambda I)(v) = 0$. Por (b), $T - \lambda I$ é normal, portanto, por (a), temos $(T - \lambda I)^*(v) = 0$. Logo $(T^* - \bar{\lambda} I)(v) = 0$; e, assim, $T^*(v) = \bar{\lambda} v$.

(d) Mostramos que $\lambda_1 \langle v, w \rangle = \lambda_2 \langle v, w \rangle$: como segue.

$$\lambda_1 \langle v, w \rangle = \langle \lambda_1 v, w \rangle = \langle T(v), w \rangle = \langle v, T^*(w) \rangle = \langle v, \bar{\lambda}_2 w \rangle = \lambda_2 \langle v, w \rangle$$

Como $\lambda_1 \neq \lambda_2$; obtemos $\langle v, w \rangle = 0$.

13.19 Demonstre o Teorema 13.13. Seja T um operador normal de um espaço com produto interno complexo V de dimensão finita. Então existe uma base ortonormal de V constituída por autovetores de T. (Assim, T pode ser representado por uma matriz diagonal em relação a alguma base ortonormal.)

A demonstração é por indução na dimensão n de V. Se dim $V = 1$, a afirmação do teorema é válida trivialmente. Suponha, então, que dim $V = n > 1$ e que o teorema seja válido para operadores normais de espaços complexos de dimensão $n - 1$. Como V é um espaço vetorial complexo, T tem, ao menos, um autovalor e, portanto, um autovetor não nulo v. Seja W o subespaço de V gerado por v e seja u_1 um vetor unitário de W.
Como v é um autovetor de T, o subespaço W é invariante por T. Contudo, v também é um autovetor de T^*, pelo Problema 13.18; logo, W também é invariante por T^*. Pelo Problema 13.18, $W^\perp$ é invariante por $T^{**} = T$. O resto da demonstração é idêntico à última parte da demonstração do Teorema 13.11 (Problema 13.14).

13.20 Demonstre o Teorema 13.14. Seja T um operador qualquer de um espaço com produto interno complexo V de dimensão finita. Então T pode ser representado por uma matriz triangular em relação a alguma base ortonormal de V.

A demonstração é por indução na dimensão n de V. Se dim $V = 1$, a afirmação do teorema é válida trivialmente. Suponha, então, que dim $V = n > 1$ e que o teorema seja válido para operadores de espaços complexos de dimensão $n - 1$. Como V é um espaço vetorial complexo, T tem, ao menos, um autovalor e, portanto, um autovetor não nulo v. Seja W o subespaço de V gerado por v e seja u_1 um vetor unitário de W. Então u_1 é um autovetor de T e, digamos, $T(u_1) = a_{11}u_1$.
Pelo Teorema 7.4, $V = W \oplus W^\perp$. Denotemos por E a projeção ortogonal de V sobre $W^\perp$. Claramente, $W^\perp$ é invariante pelo operador ET. Por indução, existe uma base ortonormal $\{u_2, ..., u_n\}$ de $W^\perp$ tal que, para $i = 2, ..., n$,

$$ET(u_i) = a_{i2}u_2 +_{i3} u_3 + \cdots + a_{ii}u_i$$

(Observe que $\{u_1, u_2, ..., u_n\}$ é uma base ortonormal de V.) Como E é a projeção ortogonal de V sobre $W^\perp$; decorre que

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{ii}u_i$$

para $i = 2, ..., n$. Isso, com $T(u_1) = a_{11}u_1$, dá o resultado procurado.

Problemas variados

13.21 Demonstre o Teorema 13.10b. Seja P um operador. As afirmações seguintes são equivalentes.

(i) $P = T^2$, para algum operador autoadjunto T.

(ii) P é não negativo, isto é, $P = S^*S$, para algum operador S.

(iii) P é autoadjunto e $\langle P(u), u \rangle \geq 0$, para cada $u \in V$ de V.

Suponha que (i) valha, ou seja, que $P = T^2$, com $T = T^*$. Então $P = TT = T^*T$ e, portanto, (i) implica (ii). Suponha, agora, que (ii) valha. Então $P^* = (S^*S)^* = S^*S^{**} = S^*S = P$ e, portanto, P é autoadjunto. Além disso,

$$\langle P(u), u \rangle = \langle S^*S(u), u \rangle = \langle S(u), S(u) \rangle \geq 0$$

Assim, (ii) implica (iii) e resta provar que (iii) implica (i).
Suponha que (iii) valha. Como P é autoadjunto, existe uma base ortonormal $\{u_1, ..., u_n\}$ consistindo em autovetores de P; digamos que $P(u_i) = \lambda_i u_i$. Pelo Teorema 13.14, os λ_i são reais. Usando (iii), mostremos que os λ_i não são negativos. Para cada i, temos

$$0 \leq \langle P(u_i), u_i \rangle = \langle \lambda_i u_i, u_i \rangle = \lambda_i \langle u_i, u_i \rangle$$

Assim, $\langle u_i, u_i \rangle \geq 0$, o que força $\lambda_i \geq 0$, como queríamos. Em vista disso, $\sqrt{\lambda_i}$ é um número real. Seja T o operador linear definido por

$$T(u_i) = \sqrt{\lambda_i} u_i, \qquad \text{com } i = 1, ..., n$$

Como T é representado por uma matriz diagonal real na base ortonormal $\{u_i\}$, decorre que T é autoadjunto. Além disso, para cada i,

$$T^2(u_i) = T(\sqrt{\lambda_i}u_i) = \sqrt{\lambda_i}T(i_i) = \sqrt{\lambda_i}\sqrt{\lambda_i}u_i = \lambda_i u_i = P(u_i)$$

Como os valores de T^2 e P coincidem numa base de V, resulta $T^2 = P$. O teorema está demonstrado.

OBSERVAÇÃO O operador T da demonstração é o único operador positivo tal que $P = T^2$, denominado *raiz quadrada positiva* de P.

13.22 Mostre que todo operador T é a soma de um operador autoadjunto e de um operador antiadjunto.

Tomamos $S = \frac{1}{2}(T + T^*)$ e $U = \frac{1}{2}(T - T^*)$. Então $T = S + U$, com

$$S^* = [\tfrac{1}{2}(T + T^*)]^* = \tfrac{1}{2}(T^* + T^{**}) = \tfrac{1}{2}(T^* + T) = S$$

e

$$U^* = [\tfrac{1}{2}(T - T^*)]^* = \tfrac{1}{2}(T^* - T) = -\tfrac{1}{2}(T - T^*) = -U$$

Assim, S é autoadjunto e T é antiadjunto.

13.23 Prove. Seja T um operador linear qualquer de um espaço com produto interno V de dimensão finita. Então T é um produto de um operador unitário (ortogonal) U e um único operador positivo P, ou seja, $T = UP$. Além disso, se T for invertível, então U também é determinado de maneira única.

Pelo Teorema 13.10, T^*T é um operador não negativo e, portanto, existe um (único) operador não negativo P tal que $P^2 = T^*T$ (Problema 13.43). Observe que

$$\|P(v)\|^2 = \langle P(v), P(v)\rangle = \langle P^2(v), v\rangle = \langle T^*T(v), v\rangle = \langle T(v), T(v)\rangle = \|T(v)\|^2 \tag{1}$$

Consideremos agora, separadamente, os casos de T ser ou não invertível.

Se T é invertível, tomamos $\hat{U} = PT^{-1}$ e mostramos que $\hat{U}$ é unitário, como segue.

$$\hat{U}^* = (PT^{-1})^* = T^{-1*}P^* = (T^*)^{-1}P \qquad \text{e} \qquad \hat{U}^*\hat{U} = (T^*)^{-1}PPT^{-1} = (T^*)^{-1}T^*TT^{-1} = I$$

Assim, $\hat{U}$ é unitário. Agora tomamos $U = \hat{U}^{-1}$. Então U é também unitário e $T = UP$, como queríamos.

Para demonstrar a unicidade, supomos que $T = U_0P_0$, com U_0 unitário e P_0 não negativo. Então

$$T^*T = P_0^*U_0^*U_0P_0 = P_0IP_0 = P_0^2$$

Mas a raiz quadrada positiva de T^*T é única (Problema 13.43), de modo que $P_0 = P$. (Observe que a invertibilidade de T não foi usada na demonstração da unicidade de P.) Agora, sendo T invertível, (1) garante que P também é invertível. Multiplicando $U_0P = UP$ à direita por P^{-1}, resulta $U_0 = U$. Assim, também U é único se T for invertível.

Suponha, agora, que T não seja invertível. Seja W a imagem de P, isto é, $W = \text{Im}\, P$. Definimos $U_1 : W \to V$ por

$$U_1(w) = T(v), \qquad \text{onde} \qquad P(v) = w \tag{2}$$

Precisamos mostrar que U_1 está bem definido, isto é, que $P(v) = P(v')$ implica $T(v) = T(v')$. Isso segue do fato de que $P(v - v') = 0$ é equivalente a $\|P(v - v')\| = 0$, o que força $\|T(v - v')\| = 0$, por (1). Assim, U_1 é um operador bem definido. A seguir, definimos $U_2 : W \to V$. Observe que, por (1), P e T têm o mesmo núcleo. Logo, as imagens de P e T têm a mesma dimensão, isto é, dim (Im P) = dim W = dim (Im T). Consequentemente, $W^\perp$ e $(\text{Im}\ T)^\perp$ também têm a mesma dimensão. Seja U_2 um isomorfismo qualquer de $W^\perp$ sobre $(\text{Im}\ T)^\perp$.

A seguir, tomamos $U = U_1 \oplus U_2$. [Aqui, definimos U como segue. Se $v \in V$ e $v = w + w'$, com $w \in W$, $w' \in W^\perp$ então $U(v) = U_1(w) + U_2(w')$.] Ocorre que U é linear (Problema 13.69) e, se $v \in V$ e $P(v) = w$, então, por (2),

$$T(v) = U_1(w) = U(w) = UP(v)$$

Assim, $T = UP$, como queríamos.

Resta mostrar que U é unitário. Todo vetor $x \in V$ pode ser escrito da forma $x = P(v) + w'$, onde $w' \in W^\perp$. Então $U(x) = UP(v) + U_2(w') = T(v) + U_2(w')$, com $\langle T(v), U_2(w')\rangle = 0$, por definição de U_2. Também $\langle T(v), T(v)\rangle = \langle P(v), P(v)\rangle$, por (1). Assim,

$$\begin{aligned}\langle U(x), U(x)\rangle &= \langle T(v) + U_2(w'),\ \ T(v) + U_2(w')\rangle = \langle T(v), T(v)\rangle + \langle U_2(w'), U_2(w')\rangle \\ &= \langle P(v), P(v)\rangle + \langle w', w'\rangle = \langle P(v) + w',\ \ P(v) + w'\rangle = \langle x, x\rangle\end{aligned}$$

[Também usamos o fato de que $\langle P(v), w'\rangle = 0$.] Assim, U é unitário e o teorema está demonstrado.

13.24 Seja V o espaço vetorial dos polinômios sobre **R** com o produto interno definido por

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt$$

Dê um exemplo de um funcional linear ϕ de V para o qual não valha o Teorema 13.3, ou seja, para o qual não exista polinômio $h(t)$ algum tal que $\phi(f) = \langle f, h \rangle$, para cada $f \in V$.

Seja $\phi : V \to \mathbf{R}$ definido por $\phi(f) = f(0)$, ou seja, ϕ calcula $f(t)$ em 0 e, portanto, leva $f(t)$ em seu termo constante. Suponha que $h(t)$ seja um polinômio tal que

$$\phi(f) = f(0) = \int_0^1 f(t)h(t)\, dt \tag{1}$$

para cada polinômio $f(t)$. Observe que ϕ leva o polinômio $tf(t)$ em 0, portanto, por (1),

$$\int_0^1 tf(t)h(t)\, dt = 0 \tag{2}$$

para cada polinômio $f(t)$. Em particular, (2) deve valer para o polinômio $f(t) = th(t)$, ou seja,

$$\int_0^1 t^2h^2(t)\, dt = 0$$

Essa integral força $h(t)$ a ser o polinômio nulo, portanto, $\phi(f) = \langle f, h \rangle = \langle f, 0 \rangle = 0$ para cada polinômio $f(t)$. Isso contradiz o fato de que ϕ não é o funcional nulo. Assim, tal polinômio $h(t)$ não existe.

Problemas Complementares

Operadores adjuntos

13.25 Encontre a adjunta da matriz dada.

(a) $A = \begin{bmatrix} 5-2i & 3+7i \\ 4-6i & 8+3i \end{bmatrix}$, (b) $B = \begin{bmatrix} 3 & 5i \\ i & -2i \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$

13.26 Seja $T : \mathbf{R}^3 \to \mathbf{R}^3$ definido por $T(x, y, z) = (x + 2y,\ \ 3x - 4z,\ \ y)$. Encontre $T^*(x, y, z)$.

13.27 Seja $T : \mathbf{C}^3 \to \mathbf{C}^3$ definido por $T(x, y, z) = [ix + (2 + 3i)y,\ \ 3x + (3 - i)z,\ \ (2 - 5i)y + iz]$. Encontre $T^*(x, y, z)$.

13.28 Para cada funcional linear ϕ de V dado, encontre $u \in V$ tal que $\phi(v) = \langle v, u \rangle$, para cada $v \in V$.

(a) $\phi : \mathbf{R}^3 \to \mathbf{R}$ definido por $\phi(x, y, z) = x + 2y - 3z$.

(b) $\phi : \mathbf{C}^3 \to \mathbf{C}$, definido por $\phi(x, y, z) = ix + (2 + 3i)y + (1 - 2i)z$.

13.29 Seja T um operador de um espaço V de dimensão finita. Prove que a imagem de T^* é o complemento ortogonal do núcleo de T, ou seja, $\mathrm{Im}\ T^* = (\mathrm{Nuc}\ T)^\perp$. Assim, $\mathrm{pos}(T) = \mathrm{pos}(T^*)$.

13.30 Mostre que $T^*T = 0$ implica $T = 0$.

13.31 Seja V o espaço vetorial dos polinômios sobre **R** com o produto interno definido por $\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt$. Seja **D** o operador derivada de V, ou seja, $\mathbf{D}(f) = df/dt$. Mostre que não existe operador $\mathbf{D}^*$ de V tal que $\langle \mathbf{D}(f), g \rangle = \langle f, \mathbf{D}^*(g) \rangle$, para quaisquer $f, g \in V$. Assim, **D** não tem adjunto.

Operadores e matrizes unitários e ortogonais

13.32 Encontre uma matriz unitária (ortogonal) cuja primeira linha seja
(a) $(2/\sqrt{13},\ 3/\sqrt{13})$, (b) um múltiplo de (1, 1, –*i*), (c) um múltiplo de (1, –*i*, 1, –*i*).

13.33 Prove que os produtos e as inversas de matrizes ortogonais são ortogonais. (Assim, as matrizes ortogonais formam um grupo perante a multiplicação, denominado *grupo ortogonal*.)

13.34 Prove que os produtos e as inversas de matrizes unitárias são unitárias. (Assim, as matrizes unitárias formam um grupo perante a multiplicação, denominado *grupo unitário*.)

13.35 Mostre que, se uma matriz ortogonal (unitária) é triangular, então ela é diagonal.

13.36 Lembre que as matrizes complexas A e B são unitariamente equivalentes se existir uma matriz unitária P tal que $B = P^*AP$. Mostre que essa equivalência é uma relação de equivalência.

13.37 Lembre que as matrizes reais A e B são ortogonalmente equivalentes se existir uma matriz ortogonal P tal que $B = P^TAP$. Mostre que essa equivalência é uma relação de equivalência.

13.38 Seja W um subespaço de V. Dado qualquer $v \in V$, escreva $v = w + w'$, com $w \in W$, $w' \in W^\perp$. (Uma tal soma é única porque $V = W \oplus W^\perp$.) Seja $T\colon V \to V$ definido por $T(v) = w - w'$. Mostre que T é um operador unitário autoadjunto de V.

13.39 Sejam V um espaço com produto interno e $U\colon V \to V$ uma aplicação (não necessariamente linear) sobrejetora que preserva o produto interno, ou seja, $\langle U(v), U(w)\rangle = \langle u, w\rangle$, para quaisquer $v, w \in V$. Prove que U é linear e, portanto, unitário.

Operadores não negativos e positivos

13.40 Mostre que a soma de dois operadores não negativos (positivos) é um operador não negativo (positivo).

13.41 Seja T um operador linear de V e defina $f\colon V \times V \to K$ por $f(u, v) = \langle T(u), v\rangle$. Mostre que f é um produto interno de V se, e só se, T é positivo.

13.42 Seja E uma projeção ortogonal sobre algum subespaço W de V. Prove que $kI + E$ é um operador não negativo (positivo) se $k \geq 0$ ($k > 0$).

13.43 Considere o operador T definido por $T(u_i) = \sqrt{\lambda_i}u_i, i = 1, \ldots, n$, na prova do Teorema 13.10b (Problema 13.21). Mostre que T é não negativo e que é o único operador não negativo tal que $T^2 = P$.

13.44 Suponha que P seja não negativo e unitário. Prove que $P = I$.

13.45 Decida se a matriz dada é não negativa ou positiva.
(a) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$, (c) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, (d) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, (e) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, (f) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

13.46 Prove que uma matriz 2×2 complexa $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ é não negativa se, e só se, (i) $A = A^*$ e (ii) a, d e $|A| = ad - bc$ são números reais não negativos.

13.47 Prove que uma matriz diagonal A é não negativa (positiva) se, e só se, cada entrada diagonal é um número real não negativo (positivo).

Matrizes autoadjuntas e simétricas

13.48 Dado qualquer operador T, mostre que $T + T^*$ é autoadjunto e que $T - T^*$ é antiadjunto.

13.49 Seja T autoadjunto. Mostre que $T^2(v) = 0$ implica $T(v) = 0$. Use isso para provar que também $T^n(v) = 0$ implica $T(v) = 0$, com $n > 0$.

13.50 Seja V um espaço com produto interno complexo. Suponha que $\langle T(v), v\rangle$ seja real, para cada $v \in V$. Mostre que T é autoadjunto.

13.51 Sejam T_1 e T_2 operadores autoadjuntos. Mostre que T_1T_2 é autoadjunto se, e só se, T_1 e T_2 comutam, isto é, $T_1T_2 = T_2T_1$.

13.52 Para cada matriz simétrica A dada, encontre uma matriz ortogonal P e uma matriz diagonal D tais que $D = P^TAP$.

(a) $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$, (b) $A = \begin{bmatrix} 5 & 4 \\ 4 & -1 \end{bmatrix}$, (c) $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$

13.53 Encontre uma mudança de coordenadas ortogonal $X = PX'$ que diagonalize cada forma quadrática dada e encontre a correspondente forma quadrática diagonal $q(x')$.

(a) $q(x, y) = 2x^2 - 6xy + 10y^2$, (b) $q(x, y) = x^2 + 8xy - 5y^2$

(b) $q(x, y, z) = 2x^2 - 4xy + 5y^2 + 2xz - 4yz + 2z^2$

Operadores e matrizes normais

13.54 Seja $A = \begin{bmatrix} 2 & i \\ i & 2 \end{bmatrix}$. Verifique que A é normal. Encontre uma matriz unitária P tal que P^*AP seja diagonal. Encontre P^*AP.

13.55 Mostre que uma matriz triangular é normal se, e só se, é diagonal.

13.56 Prove que se T é um operador normal de V, então $\|T(v)\| = \|T^*(v)\|$, para cada $v \in V$. Prove que a recíproca vale em espaços com produto interno complexos.

13.57 Mostre que operadores autoadjuntos, antiadjuntos e unitários (ortogonais) são normais.

13.58 Seja T um operador normal. Prove as afirmações dadas.

(a) T é autoadjunto se, e só se, os autovalores de T são números reais.

(b) T é unitário se, e só se, os autovalores de T têm valor absoluto 1.

(c) T é não negativo se, e só se, os autovalores de T são números reais não negativos.

13.59 Seja T um operador normal. Mostre que T e T^* têm o mesmo núcleo e a mesma imagem.

13.60 Sejam T_1 e T_2 operadores normais que comutam. Mostre que $T_1 + T_2$ e T_1T_2 também são normais.

13.61 Seja T_1 um operador normal que comuta com um operador T_2. Mostre que T_1 também comuta com o operador T_2^*.

13.62 Demonstre a afirmação seguinte. Sejam T_1 e T_2 operadores normais de um espaço com produto interno complexo V de dimensão finita. Então existe uma base ortonormal de V consistindo de autovetores de T_1 e T_2. (Assim, T_1 e T_2 são simultaneamente diagonalizáveis.)

Isomorfismos de espaços com produto interno

13.63 Seja $S = \{u_1, ..., u_n\}$ uma base ortonormal de um espaço com produto interno V sobre K. Mostre que a aplicação $v \mapsto [v]_S$ é um isomorfismo (de espaços com produto interno) de V sobre K^n. (Aqui, $[v]_S$ denota o vetor de coordenadas de v na base S.)

13.64 Mostre que dois espaços com produto interno V e W sobre K são isomorfos se, e só se, V e W têm a mesma dimensão.

13.65 Sejam $\{u_1, ..., u_n\}$ e $\{u'_1, \ldots, u'_n\}$ bases ortonormais de V e W, respectivamente. Seja $T:V \to W$ a transformação linear definida por $T(u_i) = u'_i$, para cada i. Mostre que T é um isomorfismo.

13.66 Seja V um espaço com produto interno. Vimos que cada $u \in V$ determina um funcional linear $\hat{u}$ do espaço dual V^* pela definição $\hat{u}(v) = \langle v, u\rangle$, com $v \in V$. (Ver o texto que precede o Teorema 13.3.) Mostre que a aplicação $u \mapsto \hat{u}$ é linear e não singular e, portanto, um isomorfismo de V sobre V^*.

Problemas variados

13.67 Seja $\{u_1, ..., u_n\}$ uma base ortonormal de V. Prove as afirmações dadas.
(a) $\langle a_1u_1 + a_2u_2 + \cdots + a_nu_n,\ b_1u_1 + b_2u_2 + \cdots + b_nu_n\rangle = a_1\bar{b}_1 + a_2\bar{b}_2 + \cdots + \bar{a}_n\bar{b}_n$
(b) Seja $A = [a_{ij}]$ a representação matricial de $T\colon V \to V$ na base $\{u_i\}$. Então $a_{ij} = \langle T(u_i), u_j\rangle$.

13.68 Mostre que existe uma base ortonormal $\{u_1, ..., u_n\}$ de V consistindo de autovetores de T se, e só se, existem projeções ortogonais $E_1, ..., E_r$ e escalares $\lambda_1, \ldots, \lambda_r$ tais que

(i) $T = \lambda_1E_1 + \cdots + \lambda_rE_r$, (ii) $E_1 + \cdots + E_r = I$, (iii) $E_iE_j = 0$ com $i \neq j$

13.69 Suponha que $V = U \oplus W$ e que $T_1:U \to V$ e $T_2:W \to V$ sejam operadores lineares. Mostre que $T = T_1 \oplus T_2$ também é um operador linear. Aqui, T é definido como segue. Se $v \in V$, com $v = u + w$ e $u \in U,\ w \in W$,, então

$$T(v) = T_1(u) + T_2(w)$$

Respostas dos Problemas Complementares

Notação: $[R_1;\ R_2;\ \ldots;\ R_n]$ denota uma matriz de linhas $R_1, R_2, ..., R_n$

13.25 (a) $[5 + 2i,\ 4 + 6i;\ 3 - 7i,\ 8 - 3i]$, (b) $[3, -i;\ -5i, 2i]$, (c) $[1, 2;\ 1, 3]$

13.26 $T^*(x, y, z) = (x + 3y,\ 2x + z,\ -4y)$

13.27 $T^*(x, y, z) = [-ix + 3y,\ (2 - 3i)x + (2 + 5i)z,\ (3 + i)y - iz]$

13.28 (a) $u = (1, 2, -3)$, (b) $u = (-i,\ 2 - 3i,\ 1 + 2i)$

13.32 (a) $(1/\sqrt{13})[2, 3;\ 3, -2]$, (b) $(1/\sqrt{3})[1,\ 1 - i;\ 1 + i,\ -1]$,
(c) $\frac{1}{2}[1,\ -i,\ 1 - i;\ \sqrt{2}i,\ -\sqrt{2},\ 0;\ 1,\ -i,\ -1 + i]$

13.45 Somente (a) e (e) são não negativas; somente (e) é postiva.

13.52 (a e b) $P = (1/\sqrt{5})[2, -1;\ 1, 2]$, (c) $P = (1/\sqrt{10})[3, -1;\ 1, 3]$
(a) $D = [2, 0;\ 0, -3]$, (b) $D = [7, 0;\ 0, -3]$, (c) $D = [8, 0;\ 0, -2]$

13.53 (a) $x = (3x' - y')/\sqrt{10}$, $y = (x' + 3y')/\sqrt{10}$; (b) $x = (2x' - y')/\sqrt{5}$, $y = (x' + 2y')/\sqrt{5}$;
(c) $x = x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6}$, $y = x'/\sqrt{3} - 2z'/\sqrt{6}$, $z = x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6}$;
(a) $q(x') = \mathrm{diag}(1, 11)$; (b) $q(x') = \mathrm{diag}(3, -7)$; (c) $q(x') = \mathrm{diag}(1, 17)$

13.54 (a) $P = (1/\sqrt{2})[1, -1;\ 1, 1]$, $P^*AP = \mathrm{diag}(2 + i,\ 2 - i)$

Produtos Multilineares

A.1 INTRODUÇÃO

O material deste apêndice é muito mais abstrato do que o apresentado até aqui. Em vista disso, omitimos muitas demonstrações. Também motivamos este material com a observação seguinte.

Seja *S* uma base de um espaço vetorial *V*. O Teorema 5.2 pode ser reescrito como segue.

Teorema 5.2 Seja $g : S \to V$ a aplicação inclusão de *S* em *V*. Então, dado qualquer espaço vetorial *U* e dada qualquer aplicação $f : S \to U$, existe uma única transformação linear $f^* : V \to U$ tal que $f = f^* \circ g$.

Uma outra maneira de enunciar o fato de que $f = f^* \cdot g$ é dizer que o diagrama na Figura A-1(a) é *comutativo*.

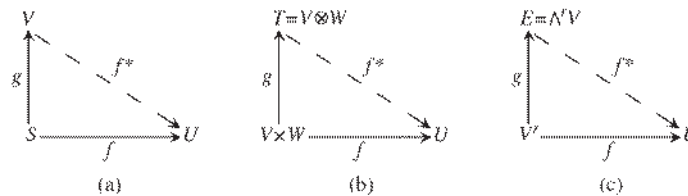



Figura A-1

A.2 APLICAÇÕES BILINEARES E PRODUTOS TENSORIAIS

Sejam *U*, *V*, *W* espaços vetoriais sobre um corpo *K*. Dizemos que uma aplicação

$$f : V \times W \to U$$

é *bilinear* se, para cada $v \in V$, a aplicação $f_v : W \to U$, definida por $f_v(w) = f(v, w)$, é linear e se, para cada $w \in W$, a aplicação $f_w : V \to U$, definida por $f_w(v) = f(v, w)$, também é linear.

Assim, *f* é linear em cada uma de suas duas variáveis. Observe que *f* é análoga a uma forma bilinear, exceto que os valores da aplicação *f* estão num espaço vetorial *U* em vez do corpo *K*.

DEFINIÇÃO Sejam *V* e *W* espaços vetoriais sobre o mesmo corpo *K*. O *produto tensorial* de *V* e *W* é um espaço vetorial *T* sobre *K* junto com uma aplicação bilinear $g : V \times W \to T$, denotada por $g(v, w) = v \otimes w$, com a propriedade que indicamos por (*), como segue. Dados qualquer espaço vetorial *U* sobre *K* e qualquer aplicação bilinear $f : V \times W \to U$, existe uma única transformação linear $f^* : T \to U$ tal que $f^* \circ g = f$.

O produto tensorial (*T*, *g*) [ou, simplesmente, *T*, quando *g* estiver subentendido] de *V* e *W* é denotado por $V \otimes W$ e o elemento $v \otimes w$ é denominado *tensor* de v e w.

Outra maneira de enunciar a condição (*) é dizer que o diagrama da Figura A-1(b) é comutativo. A existência de uma tal transformação linear única f^* é denominado "Princípio da Aplicação Universal". Conforme ilustra a Figura A-1(b), a condição (*) também diz que qualquer aplicação bilinear $f: V \times W \to U$ "fatora pelo" produto tensorial $T = V \otimes W$. A unicidade em (*) implica que a imagem de g gera T, ou seja, ger($\{v \otimes w\}$) = T.

Teorema A.1 **(Unicidade do Produto Tensorial)** Sejam (T, g) e (T', g') produtos tensoriais de V e W. Então existe um único isomorfismo $h: T \to T'$ tal que $hg = g'$.

Demonstração. Como T é um produto tensorial e $g': V \otimes W \to T'$ é bilinear, existe uma única transformação linear $h: T \to T'$ tal que $hg = g'$. Analogamente, como T' é um produto tensorial e $g: V \otimes W \to T'$ é bilinear, existe uma única transformação linear $h': T' \to T$ tal que $h'g' = g$. Usando $hg = g'$, obtemos $h'hg = g$. Além disso, como T é um produto tensorial e $g: V \otimes W \to T$ é bilinear, existe uma única transformação linear $h^*: T \to T$ tal que $h^*g = g$. No entanto, $1_T g = g$ e, portanto, $h'h = h^* = 1_T$. Analogamente, $hh' = 1_{T'}$. Assim, h é um isomorfismo de T sobre T'.

Teorema A.2 **(Existência do Produto Tensorial)** Existe o produto tensorial $T = V \otimes W$ de espaços vetoriais V e W sobre K. Sejam $\{v_1, ..., v_m\}$ uma base de V e $\{w_1, ..., w_n\}$ uma base de W. Então os mn vetores

$$v_i \otimes w_i \quad (i = 1, \ldots, m; j = 1, \ldots, n)$$

formam uma base de T. Assim, dim $T = mn$ = (dim V)(dim W).

Esboço de Demonstração. Sejam $\{v_1, ..., v_m\}$ uma base de V e $\{w_1, ..., w_n\}$ uma base de W. Considere os mn símbolos $\{t_{ij} | i = i, \ldots, m, j = 1, \ldots, n\}$. Seja T o espaço vetorial gerado pelos t_{ij}, ou seja, T consiste em todas as combinações lineares dos t_{ij} com coeficientes em K. (Ver Problema 4.137.)

Sejam $v \in V$ e $w \in W$. dados, digamos,

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m \qquad \text{e} \qquad w = b_1 w_1 + b_2 w_2 + \cdots + b_m w_m$$

Seja $g: V \times W \to T$ definida por

$$g(v, w) = \sum_i \sum_j a_i b_j t_{ij}$$

Então g é bilinear (demonstre).

Seja $f: V \times W \to U$ bilinear. Como os t_{ij} formam uma base de T, o Teorema 5.2 (enunciado no início da seção) garante que existe uma única transformação linear $f^*: T \to U$ tal que $f^*(t_{ij}) = f(v_i, w_j)$. Então, para $v = \sum_i a_i v_i$ e $w = \sum_j b_j w_j$, temos

$$f(v, w) = f\left(\sum_i a_i v_i, \sum_j b_j w_j\right) = \sum_i \sum_j a_i b_j f(v_i, w_j) = \sum_i \sum_j a_i b_j t_{ij} = f^*(g(v, w)).$$

Logo, $f = f^* g$, onde f^* é a transformação linear exigida na definição. Assim, T é um produto tensorial.

Sejam, agora, $\{v'_1, \ldots, v'_m\}$ uma base qualquer de V e $\{w'_1, \ldots, w'_m\}$ uma de W. Sejam $v \in V$ e $w \in W$ dados, digamos,

$$v = a'_1 v'_1 + \cdots + a'_m v'_m \qquad \text{e} \qquad w = b'_1 w'_1 + \cdots + b'_m w'_m$$

Então

$$v \otimes w = g(v, w) = \sum_i \sum_j a'_i b'_i g(v'_i, w'_i) = \sum_i \sum_j a'_i b'_j (v'_i \otimes w'_j)$$

Assim, os elementos $v'_i \otimes w'_j$ geram T. Há mn desses elementos, que não podem ser linearmente dependentes porque $\{t_{ij}\}$ é uma base de T e, portanto, dim $T = mn$. Assim, os $v'_i \otimes w'_j$ formam uma base de T.

Vejamos dois exemplos concretos de produtos tensoriais.

Exemplo A.1 Sejam V o espaço vetorial dos polinômios $\mathbf{P}_{r-1}(x)$ e W o espaço vetorial dos polinômios $\mathbf{P}_{s-1}(y)$. Assim os elementos a seguir formam bases de V e W, respectivamente.

$$1,\ x,\ x^2,\ \ldots,\ x^{r-1} \quad \text{e} \quad 1,\ y,\ y^2,\ \ldots,\ y^{s-1}$$

Em particular, dim $V = r$ e dim $W = s$. Seja T o espaço vetorial dos polinômios nas variáveis x e y com base

$$\{x^i y^j\}, \quad \text{com } i = 0, 1, \ldots, r-1;\ j = 0,\ 1,\ \ldots,\ s-1$$

Então T é o produto tensorial $V \otimes W$ pela aplicação

$$x^i \otimes y^j = x^i y^i$$

Por exemplo, se $v = 2 - 5x + 3x^3$ e $w = 7y + 4y^2$, então

$$v \otimes w = 14y + 8y^2 - 35xy - 20xy^2 + 21x^3y + 12x^3y^2$$

Observe que dim $T = rs =$ (dim V)(dim W).

Exemplo A.2 Sejam V o espaço vetorial das matrizes $m \times n$ sobre um corpo K e W o das matrizes $p \times q$ sobre K. Suponha que $A = [a_{ij}]$ pertença a V e B a W. Seja T o espaço vetorial das matrizes $mp \times nq$ sobre K. Então T é o produto tensorial de V e W, em que $A \otimes B$ é a matriz em blocos

$$A \otimes B = \left[a_{ij}B\right] = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

Por exemplo, se $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, então

$$A \otimes B = \begin{bmatrix} 1 & 2 & 3 & 2 & 4 & 6 \\ 4 & 5 & 6 & 8 & 10 & 12 \\ 3 & 6 & 9 & 4 & 8 & 12 \\ 12 & 15 & 18 & 16 & 20 & 24 \end{bmatrix}$$

Isomorfismos de produtos tensoriais

Inicialmente, observamos que o produto tensorial é associativo de maneira canônica, como segue.

Teorema A.3 Sejam U, V, W espaços vetoriais sobre um corpo K. Então existe um único isomorfismo

$$(U \otimes V) \otimes W \to U \otimes (V \otimes W)$$

tal que, para quaisquer $u \in U,\ v \in V,\ w \in W$,

$$(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$$

Em vista disso, podemos omitir os parênteses quando tensoriamos um número finito qualquer de fatores. Mais precisamente, dados espaços vetoriais V_1, V_2, ..., V_m sobre um corpo K, podemos definir, sem ambiguidade, o produto tensorial

$$V_1 \otimes V_2 \otimes \ldots \otimes V_m$$

e, dados vetores v_j em V_j, podemos formar, sem ambiguidade, o produto tensorial

$$v_1 \otimes v_2 \otimes \ldots \otimes v_m$$

Além disso, dado um espaço vetorial V sobre K, podemos definir, sem ambiguidade, o produto tensorial

$$\otimes^r V = V \otimes V \otimes \ldots \otimes V \text{ (de } r \text{ fatores)}$$

Também temos o isomorfismo canônico

$$(\otimes^r V) \otimes (\otimes^s V) \to \otimes^{r+s} V$$

Finalmente, vendo K como um espaço vetorial sobre si mesmo, temos o isomorfismo canônico

$$K \otimes V \to V$$

em que definimos $a \otimes v = av$.

A.3 APLICAÇÕES MULTILINEARES ALTERNADAS

Seja $f: V^r \to U$, em que V e U são espaços vetoriais sobre K. [Lembre que $V^r = V \times V \times \ldots \times V$, com r fatores.]

(1) Dizemos que f é uma *aplicação multilinear* ou *r-linear* se $f(v_1, \ldots, v_r)$ é linear como função de cada v_j separadamente, considerando os demais v_i fixados. Isto é,

$$f(\ldots, v_j + v_j', \ldots) = f(\ldots, v_j, \ldots) + f(\ldots, v_j', \ldots)$$
$$f(\ldots, kv_j, \ldots) = kf(\ldots, v_j, \ldots)$$

em que apenas a j-ésima posição varia.

(2) Dizemos que a aplicação f é *alternada* se

$$f(v_1, \ldots, v_r) = 0, \text{ sempre que } v_i = v_j, \text{ com } i \neq j.$$

Pode-se mostrar facilmente (demonstre) que se f é uma aplicação multilinear alternada de V^r, então

$$f(\ldots, v_i, \ldots, v_j, \ldots) = -f(\ldots, v_j, \ldots, v_i, \ldots)$$

Isto é, se trocamos de posição dois vetores, então o valor associado troca de sinal.

Exemplo A.3 **(Determinantes)** A função determinante $D: M \to K$ do espaço M das matrizes $n \times n$ pode ser vista como uma aplicação com n variáveis

$$D(A) = D(R_1, R_2, \ldots, R_n)$$

definida nas linhas $R_1, R_2, ..., R_n$ de A. Na Seção 8.15 vimos que, nesse contexto, D é tanto n-linear quanto alternada.

Precisamos, agora, de mais notação. Denotemos por $P = [k_1, k_2, ..., k_r]$ uma lista de r elementos de $I_n = \{1, 2, ..., n\}$ (uma r-lista). Então utilizamos a notação a seguir, em que os v_k denotam vetores e os a_{ik} denotam escalares.

$$v_P = (v_{k_1}, v_{k_2}, \ldots, v_{k_r}) \text{ e } a_P = a_{1k_1} a_{2k_2} \ldots a_{rk_r}$$

Observe que v_P é uma lista de r vetores e a_P é um produto de r escalares.

Suponha, agora, que os elementos de $P = [k_1, k_2, ..., k_r]$ sejam distintos. Então P é uma permutação σ_P de uma r-lista $J = [i_1, i_2, ..., i_r]$ em *forma padrão*, isto é, tal que $i_1 < i_2 < \cdots < i_r$. O número dessas listas padrão J de r elementos de I_n é o coeficiente binomial

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

[Na Seção 8.4 vimos que $\operatorname{sgn}(\sigma_K) = (-1)^{m_P}$, onde m_P é o número de inversões que transformam P em J.]

Seja, agora, $A = [a_{ij}]$ uma matriz $r \times n$. Dada uma r-lista ordenada J, definimos

$$D_J(A) = \begin{vmatrix} a_{1i_1} & a_{1i_2} & \ldots & a_{1i_r} \\ a_{2i_1} & a_{2i_2} & \ldots & a_{2i_r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{ri_1} & a_{ri_2} & \ldots & a_{ri_r} \end{vmatrix}$$

Isto é, $D_J(A)$ é o determinante da submatriz $r \times r$ de A cujos índices de coluna pertencem a J.

A demonstração do teorema principal desta seção (Teorema A.6) necessita de um lema relativo a "embaralhamentos".

Lema A.4 Sejam V e U espaços vetoriais sobre K e $f : V^r \to U$ uma aplicação r-linear alternada. Sejam $v_1, v_2, ..., v_n$ vetores de V e $A = [a_{ij}]$ uma matriz $r \times n$ sobre K, com $r \leq n$. Para $i = 1, 2, ..., r$, escrevemos

$$u_i = a_{i1}v_i + a_{i2}v_2 + \cdots + a_{in}v_n$$

Então

$$f(u_1, \ldots, u_r) = \sum_f D_J(A) f(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$$

em que a soma é sobre todas as r-listas $J = [i_1, i_2, ..., i_r]$ em forma padrão.

A demonstração é técnica, mas direta. A linearidade de f fornece a soma

$$f(u_1, \ldots, u_r) = \sum_K a_K f(v_K)$$

em que a soma é sobre todas as r-listas P de $\{1, ..., n\}$. Por ser alternada, temos $f(v_P) = 0$, sempre que P tiver algum inteiro repetido. Agora, a demonstração utiliza, principalmente, o fato de que quando trocamos a posição dos v_j para transformar

$$f(v_P) = f(v_{k_1}, v_{k_2}, \ldots, v_{k_r}) \quad \text{em} \quad f(v_J) = f(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$$

de tal forma que $i_1 < \cdots < i_r$, o sinal associado a a_P mudará do mesmo modo que o sinal da permutação σ_P correspondente muda quando é transformada na permutação identidade por meio de transposições.

No exemplo, ilustramos o lema com $r = 2$ e $n = 3$.

Exemplo A.4 Seja $f : V^2 \to U$ uma aplicação multilinear alternada. Sejam $v_1, v_2, v_3 \in V$ e $u, w \in V$ e suponha que

$$u = a_1v_1 + a_2v_2 + a_3v_3 \quad \text{e} \quad w = b_1v_1 + b_2v_2 + b_3v_3$$

Considere

$$f(u, w) = f(a_1v_1 + a_2v_2 + a_3v_3,\ b_1v_1 + b_2v_2 + b_3v_3)$$

Usando multilinearidade, obtemos as nove parcelas

$$\begin{aligned} f(u, w) = {} & a_1b_1 f(v_1, v_r) + a_1b_2 f(v_1, v_2) + a_1b_3 f(v_1, v_3) \\ & + a_2b_1 f(v_2, v_1) + a_2b_2 f(v_2, v_2) + a_2b_3 f(v_2, v_3) \\ & + a_3b_1 f(v_3, v_1) + a_3b_2 f(v_3, v_2) + a_3b_3 f(v_3, v_3) \end{aligned}$$

(Observe que $J = [1, 2]$, $J' = [1, 3]$ e $J'' = [2, 3]$ são as três 2-listas em forma padrão de $I_3 = \{1, 2, 3\}$.) Como f é alternada, temos $f(v_i, v_i) = 0$; logo, três dessas nove parcelas são iguais a 0. Novamente por ser f alternada, temos $f(v_i, v_j) = -f(v_j, v_i)$. Assim três dos termos podem ser transformados, de tal modo que seus índices formem uma 2-lista em forma padrão com uma única inversão. Finalmente, obtemos

$$\begin{aligned} f(u, w) &= (a_1b_2 - a_2b_1) f(v_1, v_2) + (a_1b_3 - a_3b_1) f(v_1, v_3) + (a_2b_3 - a_3b_2) f(v_2, v_3) \\ &= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} f(v_1, v_2) + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} f(v_1, v_3) + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} f(v_2, v_3) \end{aligned}$$

que é o conteúdo do Lema A.4.

A.4 PRODUTOS EXTERIORES

Temos a definição a seguir.

DEFINIÇÃO Sejam V um espaço vetorial de dimensão n sobre um corpo K e r um inteiro tal que $1 \leq r \leq n$. O *produto exterior de potência r* de V (ou, simplesmente, *produto exterior* de V, se o r estiver subentendido) é um espaço vetorial E sobre K junto com uma aplicação r-linear alternada $g: V^r \to \mathbf{E}$, denotada por $g(v_1, \ldots, v_r) = v_1 \wedge \ldots \wedge v_r$, com a propriedade que indicamos por (*), como segue. Dado qualquer espaço vetorial U sobre K e qualquer aplicação r-linear alternada $f: V^r \to U$, existe uma única transformação linear $f^*: \mathbf{E} \to U$ tal que $f^* \circ g = f$.

O produto exterior (E, g) [ou, simplesmente, E, quando g estiver subentendido] de potência r de V é denotado por $\wedge^r V$ e o elemento $v_1 \wedge \cdots \wedge v_r$ é denominado *produto exterior* dos v_i.

Uma outra maneira de enunciar a condição (*) é dizer que o diagrama da Figura A-1(c) é comutativo. Novamente, a existência de uma tal transformação linear única f^* é denominado "Princípio da Aplicação Universal". Conforme ilustra a Figura A-1(c), a condição (*) também diz que qualquer aplicação r-linear alternada $f: V^r \to U$ "fatora pelo" produto exterior $E = \wedge^r V$. Novamente, a unicidade em (*) implica que a imagem de g gera E, ou seja, $\text{ger}(\{v_1 \wedge \cdots \wedge v_r\}) = E$.

Teorema A.5 **(Unicidade do Produto Exterior)** Sejam (E, g) e (E', g') produtos exteriores de potência r de V. Então existe um único isomorfismo $h: E \to E'$ tal que $hg = g'$.

A demonstração é análoga à do Teorema A.1, usando o princípio da aplicação universal.

Teorema A.6 **(Existência do Produto Exterior)** Existe o produto exterior $E = \wedge^r V$ de um espaço vetorial V sobre K. Se $r > n$, então $E = \{0\}$. Se $r \leq n$, então $\dim E = \binom{n}{r}$. Além disso, se $\{v_1, ..., v_m\}$ é uma base de V, então os vetores

$$v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_r},$$

com $1 \leq i_1 < i_2 < \cdots < i_r \leq n$, formam uma base de E.

Vejamos um exemplo concreto de produto exterior.

Exemplo A.5 **(Produto vetorial)** Considere $V = \mathbf{R}^3$ com a base canônica $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Seja $E = \wedge^2 V$. Observe que dim $V = 3$. Assim, dim $E = 3$, com base $\mathbf{i} \wedge \mathbf{j}$, $\mathbf{i} \wedge \mathbf{k}$, $\mathbf{j} \wedge \mathbf{k}$. Identificamos E com $\mathbf{R}^3$ pela correspondência

$$\mathbf{i} = \mathbf{j} \wedge \mathbf{k},\ \mathbf{j} = \mathbf{k} \wedge \mathbf{i} = -\mathbf{i} \wedge \mathbf{k},\ \mathbf{k} = \mathbf{i} \wedge \mathbf{j}$$

Sejam u e w vetores quaisquer de $V = \mathbf{R}^3$, digamos,

$$u = (a_1, a_2, a_3) = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{e} \quad w = (b_1, b_2, b_3) = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

Então, como no Exemplo A.3,

$$u \wedge w = (a_1b_2 - a_2b_1)(\mathbf{i} \wedge \mathbf{j}) + (a_1b_3 - a_3b_1)(\mathbf{i} \wedge \mathbf{k}) + (a_2b_3 - a_3b_2)(\mathbf{j} \wedge \mathbf{k})$$

Usando a identificação, obtemos

$$\begin{aligned} u \wedge w &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k} \end{aligned}$$

Assim, esse produto exterior é precisamente o produto vetorial conhecido de $\mathbf{R}^3$ (ver Seção 1.6).

Nosso último teorema afirma que é possível "multiplicar" produtos exteriores, o que nos permite formar uma "álgebra exterior", como segue.

Teorema A.7 Sejam V um espaço vetorial sobre K e r, s inteiros positivos. Existe uma única aplicação bilinear

$$\bigwedge^r V \times \bigwedge^s V \to \bigwedge^{r+s} V$$

tal que, dados quaisquer vetores u_i, w_j em V,

$$(u_1 \wedge \cdots \wedge u_r) \times (w_1 \wedge \cdots \wedge w_s) \mapsto u_1 \wedge \cdots \wedge u_r \wedge w_1 \wedge \cdots \wedge w_s$$

Exemplo A.6 Formamos uma álgebra exterior A sobre um corpo K usando variáveis não comutativas x, y e z. Por ser uma álgebra exterior, nossas variáveis satisfazem

$$x \wedge x = 0, \quad y \wedge y = 0, \quad z \wedge z = 0, \quad \text{e} \quad y \wedge x = -x \wedge y, \quad z \wedge x = -x \wedge z, \quad z \wedge y = -y \wedge z$$

Cada elemento de A é uma combinação linear dos oito elementos

$$1, \quad x, \quad y, \quad z, \quad x \wedge y, \quad x \wedge z, \quad y \wedge z, \quad x \wedge y \wedge z$$

Multiplicamos dois "polinômios" em A usando a distributividade usual, mas também as condições dadas. Por exemplo,

$$[3 + 4y - 5x \wedge y + 6x \wedge z] \wedge [5x - 2y] = 15x - 6y - 20x \wedge y + 12x \wedge y \wedge z$$

Observe que usamos o fato de que

$$[4y] \wedge [5x] = 20y \wedge x = -20x \wedge y \quad \text{e} \quad [6x \wedge z] \wedge [-2y] = -12x \wedge z \wedge y = 12x \wedge y \wedge z$$

Estruturas Algébricas

B.1 INTRODUÇÃO

Neste apêndice definimos estruturas algébricas que ocorrem em quase todas as áreas da Matemática. Em particular, definimos *corpo*, que aparece na definição de espaço vetorial. Começamos com a definição de *grupo*, que é uma estrutura algébrica relativamente simples, com apenas uma operação, e que é usado como um bloco com o qual podemos construir muitos outros sistemas algébricos.

B.2 GRUPOS

Seja G um conjunto não vazio com uma operação binária, isto é, a cada par de elementos $a,\ b \in G$ está associado um elemento $ab \in G$. Dizemos que G é um *grupo* se os axiomas a seguir forem válidos.

[G_1] Dados quaisquer $a,\ b,\ c \in G$, temos $(ab)c = a(bc)$ (*associatividade*).
[G_2] Existe um elemento $e \in G$, denominado elemento *neutro*, ou *identidade*, tal que $ae = ea = a$, para cada $a \in G$.
[G_3] Para cada $a \in G$, existe um elemento $a^{-1} \in G$, denominado *inverso* de a, tal que $aa^{-1} = a^{-1}a = e$.

Dizemos que um grupo G é *abeliano* (ou *comutativo*) se valer a *comutatividade*, isto é, $ab = ba$ para quaisquer $a,\ b \in G$.

Quando a operação binária é denotada por justaposição, como aqui, dizemos que G é um *grupo multiplicativo*. Às vezes, quando o grupo é abeliano, denotamos a operação binária por $+$ e, nesse caso, dizemos que G é um *grupo aditivo*. Nesse caso, o elemento neutro é denotado por 0 e denominado elemento *zero*, e o elemento inverso de a é denotado $-a$ e denominado elemento *simétrico* ou *oposto* de a.

Se A e B são subconjuntos de um grupo G, escrevemos

$$AB = \{ab | a \in A,\ b \in B\} \quad \text{ou} \quad A + B = \{a + b | a \in A,\ b \in B\}$$

Também escrevemos a em vez de $\{a\}$.

Um subconjunto H de um grupo G é um *subgrupo* de G se o próprio H formar um grupo com a operação induzida de G. Se H é um subgrupo de G e $a \in G$, dizemos que o conjunto Ha é uma *classe lateral à direita* de H em G e o conjunto aH é uma *classe lateral à esquerda* de H em G.

DEFINIÇÃO Um subgrupo H de G é um *subgrupo normal* se $a^{-1}Ha \subseteq H$, para cada $a \in G$. Equivalentemente, H é normal se $aH = Ha$, para cada $a \in G$, isto é, se as classes laterais à esquerda e à direita de H coincidem.

Observe que cada subgrupo de um grupo abeliano é normal.

Teorema B.1 Seja H um subgrupo normal de G. Então as classes laterais de H em G formam um grupo com a multiplicação de classes. Esse grupo é denominado *grupo quociente* e denotado por G/H.

Exemplo B.1 O conjunto **Z** dos inteiros forma um grupo abeliano com a adição. (Observe que os inteiros pares formam um subgrupo de **Z**, mas não os ímpares.) Denotemos por H o conjunto dos múltiplos de 5, isto é, $H = \{\ldots, -10, -5, 0, 5, 10, \ldots\}$. Então H é um subgrupo (necessariamente normal) de **Z**. As classes laterais de H em **Z** são dadas a seguir.

$$\begin{aligned}
\bar{0} &= 0 + H = H = \{\ldots, -10, -5, 0, 5, 10, \ldots\} \\
\bar{1} &= 1 + H = \{\ldots, -9, -4, 1, 6, 11, \ldots\} \\
\bar{2} &= 2 + H = \{\ldots, -8, -3, 2, 7, 12, \ldots\} \\
\bar{3} &= 3 + H = \{\ldots, -7, -2, 3, 8, 13, \ldots\} \\
\bar{4} &= 4 + H = \{\ldots, -6, -1, 4, 9, 14, \ldots\}
\end{aligned}$$

Dado qualquer inteiro $n \in \mathbf{Z}$, $\bar{n} = n + H$ coincide com uma das classes listadas. Assim, pelo teorema precedente, $\mathbf{Z}/H = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ forma um grupo com a soma de classes laterais. A tabela da adição é dada a seguir.

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

Esse grupo quociente é denominado grupo dos inteiros módulo 5, frequentemente denotado por $\mathbf{Z}_5$. Analogamente, dado qualquer inteiro positivo n, construímos o grupo $\mathbf{Z}_n$ dos inteiros módulo n.

Exemplo B.2 As permutações de n símbolos (ver Seção 8.4) formam um grupo com a composição de aplicações, denominado *grupo simétrico* de grau n e denotado por $\mathbf{S}_n$. Investigamos $\mathbf{S}_3$, cujos elementos são

$$\epsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad \phi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \qquad \phi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Aqui $\begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix}$ é a permutação que transforma $1 \mapsto i$, $2 \mapsto j$, $3 \mapsto k$. A tabela de multiplicação é dada a seguir.

	ϵ	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
ϵ	ϵ	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
σ_1	σ_1	ϵ	ϕ_1	ϕ_2	σ_2	σ_3
σ_2	σ_2	ϕ_2	ϵ	ϕ_1	ϕ_3	σ_1
σ_3	σ_3	ϕ_1	ϕ_2	ϵ	σ_1	σ_2
ϕ_1	ϕ_1	σ_3	σ_1	σ_2	ϕ_2	ϵ
ϕ_2	ϕ_2	σ_2	σ_3	σ_1	ϵ	ϕ_1

(Na tabela, *ab* é o elemento na linha *a* e coluna *b*.) O conjunto $H = \{\epsilon, \sigma_1\}$ é um subgrupo de $\mathbf{S}_3$, com classes laterais à direita e à esquerda como segue.

Classes laterais à direita	Classes laterais à esquerda
$H = \{\epsilon, \sigma_1\}$	$H = \{\epsilon, \sigma_1\}$
$H_{\phi_1} = \{\phi_1, \sigma_2\}$	$\phi_2 H = \{\phi_1, \sigma_3\}$
$H_{\phi_2} = \{\phi_2, \sigma_3\}$	$\phi_2 H = \{\phi_2, \sigma_2\}$

Observe que as classes laterais à direita e à esquerda são diferentes, portanto, H não é um subgrupo normal.

Uma aplicação f de um grupo G num grupo G' é denominada *homomorfismo* (*de grupos*) se $f(ab) = f(a)f(b)$, para quaisquer $a, b \in G$. (Se f também é uma bijeção, isto é, injetora e sobre, dizemos que f é um isomorfismo e que G e G' são isomorfos.) Se $f : G \to G'$ é um homomorfismo, então o *núcleo* de f é o conjunto dos elementos de G que f leva no neutro $e' \in G'$, ou seja,

$$\text{núcleo de } f = \{a \in G \,|\, f(a) = e'\}$$

(Como sempre, o conjunto $f(G)$ é denominado *imagem* da aplicação $f : G \to G'$.) Vale o teorema seguinte.

Teorema B.2 Seja $f: G \to G$ um homomorfismo. Então o núcleo N de f é um subgrupo normal de G e o grupo quociente G/N é isomorfo à imagem de f.

Exemplo B.3 Seja G o grupo dos números reais com a adição e G' o grupo dos números reais positivos com a multiplicação. A aplicação $f : G \to G'$ definida por $f(a) = 2^a$ é um homomorfismo, pois

$$f(a+b) = 2^{a+b} = 2^a 2^b = f(a)f(b)$$

Em particular, f é uma bijeção, de modo que G e G' são isomorfos.

Exemplo B.4 Seja G o grupo dos números complexos não nulos com a multiplicação e G' o grupo dos números reais não nulos com a multiplicação. A aplicação $f : G \to G'$ definida por $f(z) = |z|$ é um homomorfismo, pois

$$f(z_1 z_2) = |z_1 z_2| = |z_1||z_2| = f(z_1)f(z_2)$$

O núcleo K de f consiste nos números complexos z do círculo unitário, isto é, tais que $|z| = 1$. Assim, G/K é isomorfo à imagem de f, ou seja, ao grupo de números reais positivos com a multiplicação.

B.3 ANÉIS, DOMÍNIOS E CORPOS

Seja R um conjunto não vazio com duas operações binárias, uma adição, denotada por $+$ e uma multiplicação, denotada por justaposição. Dizemos que R é um *anel* se os axiomas a seguir forem válidos.

[R_1] Dados quaisquer $a, b, c \in R$, temos $(a + b) + c = a + (b + c)$.
[R_2] Existe um elemento $0 \in R$, denominado elemento *zero*, ou *nulo*, tal que $a + 0 = 0 + a = a$, para cada $a \in R$.
[R_3] Para cada $a \in R$, existe um elemento $-a \in R$, denominado *oposto*, ou *simétrico*, de a, tal que $a + (-a) = (-a) + a = 0$.
[R_4] Dados quaisquer $a, b \in R$, temos $a + b = b + a$.
[R_5] Dados quaisquer $a, b, c \in R$, temos $(ab)c = a(bc)$.
[R_6] Dados quaisquer $a, b, c \in R$, temos

(i) $a(b + c) = ab + ac$ e (ii) $(b + c)a = ba + ca$.

Observe que os axiomas [R_1] a [R_4] podem ser resumidos dizendo que R é um grupo abeliano na adição.

A subtração é definida em R por $a - b = a + (-b)$.

Pode ser mostrado (Problema B.25) que $a0 = 0a = 0$, para cada $a \in R$.

Dizemos que um anel R é *comutativo* se $ab = ba$, para quaisquer $a, b \in R$. Dizemos que R é um *anel com unidade* se existir um elemento não nulo $1 \in R$ tal que $a1 = 1a = a$, para cada $a \in R$.

Um subconjunto S de um anel R é um *subanel* de R se o próprio S formar um anel com as operações induzidas de R. Observamos que S é um subanel de R se, e só se, $a, b \in S$ implica $a - b \in S$ e $ab \in S$.

Dizemos que um subconjunto I de R é um *ideal à esquerda* de R se (i) $a - b \in I$, para quaisquer $a, b \in I$ e (ii) $ra \in I$, para quaisquer $r \in R$, $a \in I$. Observe que um ideal à esquerda I de R também é um subanel de R. Analogamente, definimos *ideais à direita* e *ideais bilaterais*. Claramente, todos os ideais de anéis comutativos são bilaterais. O termo *ideal* significa ideal bilateral, salvo menção explícita em contrário.

Teorema B.3 Seja I um ideal (bilateral) de um anel R. Então as classes laterais $\{a + I \mid a \in R\}$ formam um anel com a adição e multiplicação de classes. Esse anel é denominado *anel quociente* e denotado por R/I.

Seja R, agora, um anel comutativo com unidade. Para cada $a \in R$, o conjunto $(a) = \{ra \mid r \in R\}$ é um ideal, denominado *ideal principal gerado por a*. Se todo ideal de R for principal, diremos que R é um *anel de ideais principais*.

DEFINIÇÃO Um anel comutativo com unidade R é denominado *domínio de integridade* ou, simplesmente, *domínio*, se R não tiver divisores de zero, ou seja, se $ab = 0$ implicar $a = 0$ ou $b = 0$ em R.

DEFINIÇÃO Um anel comutativo com unidade R é denominado *corpo* se todo $a \in R$ não nulo tiver um inverso multiplicativo, isto é, se existir um elemento $a^{-1} \in R$ tal que $aa^{-1} = a^{-1}a = 1$.

Um corpo é, necessariamente, um domínio, pois $ab = 0$ e $a \neq 0$ implicam

$$b = 1 \cdot b = a^{-1}ab = a^{-1} \cdot 0 = 0$$

Observamos que um corpo também pode ser visto como um anel comutativo em que os elementos não nulos formam um grupo com a multiplicação do anel.

Exemplo B.5 O conjunto $\mathbf{Z}$ dos inteiros com as operações usuais de adição e multiplicação é o exemplo clássico de domínio com unidade. Cada ideal I de $\mathbf{Z}$ é principal, ou seja, $I = (n)$ para algum inteiro n. O anel quociente $\mathbf{Z}_n = \mathbf{Z}/(n)$ é denominado *anel dos inteiros módulo n*. Se n for primo, então $\mathbf{Z}_n$ é um corpo. Por outro lado, se n não for primo, então $\mathbf{Z}_n$ tem divisores de zero. Por exemplo, no anel $\mathbf{Z}_n$, temos $\mathbf{Z}_6$, $\bar{2}\bar{3} = \bar{0}$, com $\bar{2} \neq \bar{0}$ e $\bar{3} \neq \bar{0}$.

Exemplo B.6 Os números racionais $\mathbf{Q}$ e os números reais $\mathbf{R}$ formam um corpo com as operações usuais de adição e multiplicação.

Exemplo B.7 Seja $\mathbf{C}$ o conjunto dos pares ordenados de números reais com a adição e a multiplicação definidas por

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

Então $\mathbf{C}$ satisfaz todas as propriedades requeridas de um corpo. De fato, $\mathbf{C}$ é simplesmente o corpo dos números complexos (ver Seção 1.7).

Exemplo B.8 O conjunto M de todas as matrizes 2×2 com entradas reais forma um anel comutativo com divisores de zero com as operações de adição matricial e multiplicação matricial.

Exemplo B.9 Seja R um anel qualquer. Então o conjunto $R[x]$ de todos os polinômios sobre R forma um anel com as operações usuais de adição e multiplicação polinomiais. Além disso, se R for um domínio, então $R[x]$ também será um domínio.

Seja D, agora, um domínio. Dizemos que b *divide* a em D se $a = bc$, para algum $c \in D$. Dizemos que um elemento $u \in D$ é *invertível* se u divide 1, ou seja, se u tiver inverso multiplicativo. Dizemos que um elemento $b \in D$ é *associado* de $a \in D$ se $b = ua$, para algum invertível $u \in D$. Um elemento não invertível $p \in D$ é denominado *irredutível* se $p = ab$ implica que a ou b é invertível.

Dizemos que um domínio é um *domínio de fatoração única* se todo elemento não invertível de D puder ser escrito de modo único (salvo associados e ordem) como um produto de elementos irredutíveis.

Exemplo B.10 O anel $\mathbf{Z}$ dos inteiros é o exemplo clássico de domínio de fatoração única. Os elementos invertíveis de $\mathbf{Z}$ são 1 e –1. Os únicos associados de $n \in \mathbf{Z}$ são n e $-n$. Os elementos irredutíveis de $\mathbf{Z}$ são os números primos.

Exemplo B.11 O conjunto $D = \{a + b\sqrt{13} \mid a, b \text{ inteiros}\}$ é um domínio. Os invertíveis de D são ± 1, $18 \pm 5\sqrt{13}$ e $-18 \pm 5\sqrt{13}$. Os elementos 2, $3 - \sqrt{13}$ e $-3 - \sqrt{13}$ são irredutíveis em D. Observe que

$$4 = 2 \cdot 2 = \left(3 - \sqrt{13}\right)\left(-3 - \sqrt{13}\right).$$

Assim, D não é um domínio de fatoração única. (Ver Problema B.40.)

B.4 MÓDULOS

Sejam M um grupo abeliano aditivo e R um anel com unidade. Dizemos que M é um *R-módulo* (à esquerda), ou um *módulo* (à esquerda) sobre R, se existir uma aplicação $R \times M \to M$, denotada por justaposição, tal que

[M_1] $r(m_1 + m_2) = rm_1 + rm_2$
[M_2] $(r + s)m = rm + sm$
[M_3] $(rs)m = r(sm)$
[M_4] $1m = m$

para quaisquer $r, s \in R$ e $m, m_1, m_2 \in M$.

Enfatizamos que um R-módulo é uma generalização de um espaço vetorial, já que, num módulo, permitimos que os escalares pertençam a um anel em vez de pertencerem a um corpo.

Exemplo B.12 Seja G um grupo abeliano aditivo qualquer. O grupo G é um módulo sobre o anel $\mathbf{Z}$ dos inteiros definindo

$$ng = \overbrace{g + g + \cdots + g}^{n \text{ vezes}}, \quad 0g = 0, \quad (-n)g = -ng$$

onde n é um inteiro positivo qualquer.

Exemplo B.13 Sejam R um anel e I um ideal de R. Então I pode ser visto como um módulo sobre R.

Exemplo B.14 Sejam V um espaço vetorial sobre um corpo K e $T: V \to V$ um operador linear. Fazemos de V um módulo sobre o anel $K[x]$ dos polinômios sobre K definindo $f(x)v = f(T)(v)$. O leitor deveria verificar que isto define uma multiplicação por escalar.

Seja M um R-módulo. Um subgrupo aditivo N de M é denominado *submódulo* de M se $u \in N$ e $r \in R$ implicam $ru \in N$. (Observe que, nesse caso, N é um R-módulo.)

Sejam M e M' R-módulos. Dizemos que uma aplicação $T: M \to M'$ é um *homomorfismo* (ou *R-homomorfismo* ou *R-linear*) se

(i) $T(u + v) = T(u) + T(v)$ e (ii) $T(ru) = rT(u)$

para quaisquer $u, v \in M$ e $r \in R$.

Problemas

Grupos

B.1 Decida se o sistema dado forma um grupo G.

(i) G = conjunto dos inteiros, operação de subtração.

(ii) $G = \{1, -1\}$, operação de multiplicação.

(iii) G = conjunto dos números racionais não nulos, operação de divisão.

(iv) G = conjunto das matrizes $n \times n$ não singulares, operação de multiplicação matricial.

(v) $G = \{a + bi : a, b \in \mathbf{Z}\}$, operação de adição.

B.2 Mostre que, num grupo G, valem as afirmações a seguir.

(i) O elemento neutro de G é único.

(ii) Cada $a \in G$ tem um único elemento inverso $a^{-1} \in G$.

(iii) $(a^{-1})^{-1} = a$, e $(ab)^{-1} = b^{-1}a^{-1}$.

(iv) $ab = ac$ implica $b = c$ e $ba = ca$ implica $b = c$.

B.3 Num grupo G, definimos as potências de $a \in G$ por

$$a^0 = e, \quad a^n = aa^{n-1}, \quad a^{-n} = (a^n)^{-1}, \text{onde } n \in N$$

Mostre que valem as fórmulas seguintes, para quaisquer inteiros $r, s, t \in \mathbf{Z}$.
(i) $a^r a^s = a^{r+s}$, (ii) $(a^r)^s = a^{rs}$, (iii) $(a^{r+s})^t = a^{rs+st}$.

B.4 Mostre que se G é um grupo abeliano, então $(ab)^n = a^n b^n$, para quaisquer $a, b \in G$ e qualquer inteiro $n \in \mathbf{Z}$.

B.5 Suponha que G seja um grupo tal que $(ab)^2 = a^2 b^2$, para quaisquer $a, b \in G$. Mostre que G é abeliano.

B.6 Seja H um subconjunto de um grupo G. Mostre que H é um subgrupo de G se, e só se, (i) H é não vazio e (ii) $a, b \in H$ implica $ab^{-1} \in H$.

B.7 Prove que a interseção de um número qualquer de subgrupos de G também é um subgrupo de G.

B.8 Mostre que o conjunto de todas as potências de $a \in G$ é um subgrupo de G, denominado *subgrupo cíclico* gerado por a.

B.9 Dizemos que um grupo G é *cíclico* se G for gerado por algum $a \in G$, ou seja, se $G = \{a^n : n \in Z\}$. Mostre que todo subgrupo de um grupo cíclico é cíclico.

B.10 Seja G um grupo cíclico. Mostre que G é isomorfo ao grupo $\mathbf{Z}$ dos inteiros com a adição ou ao grupo $\mathbf{Z}_n$ dos inteiros módulo n com a adição.

B.11 Seja H um subgrupo de G. Mostre que as classes laterais à direita (à esquerda) de H particionam G em conjuntos dois a dois disjuntos.

B.12 A *ordem* de um grupo G, denotada por $|G|$, é o número de elementos de G. Prove que se H é um subgrupo de um grupo finito G, então $|H|$ divide $|G|$. (Este é o Teorema de Lagrange).

B.13 Suponha que $|G| = p$, com p primo. Mostre que G é cíclico.

B.14 Sejam H e N subgrupos de G, com N normal. Mostre que (i) HN é um subgrupo de G e (ii) $H \cap N$ é um subgrupo normal de G.

B.15 Seja H um subgrupo de G com somente duas classes laterais à direita (à esquerda). Mostre que H é um subgrupo normal de G.

B.16 Demonstre o Teorema B.1 Seja H um subgrupo normal de G. Então as classes laterais de H em G formam um grupo G/H com a multiplicação de classes.

B.17 Seja G um grupo abeliano. Mostre que todo grupo quociente G/H também é abeliano.

B.18 Seja $f : G \to G'$ um homomorfismo de grupos. Prove as afirmações dadas.

(i) $f(e) = e'$, em que e e e' são as unidades de G e G', respectivamente.

(ii) $f(a^{-1}) = f(a)^{-1}$, para cada $a \in G$.

B.19 Demonstre o Teorema B.2. Seja $f : G \to G'$ um homomorfismo. Então o núcleo N de f é um subgrupo normal de G e o grupo quociente G/N é isomorfo à imagem de f.

B.20 Seja G o grupo multiplicativo dos números complexos z tais que $|z| = 1$ e seja $\mathbf{R}$ o grupo multiplicativo dos números reais. Prove que G é isomorfo a $\mathbf{R}/\mathbf{Z}$.

B.21 Fixado $g \in G$, defina $\hat{g} : G \to G$ por $\hat{g}(a) = g^{-1}ag$. Mostre que G é um isomorfismo de G sobre G.

B.22 Seja G o grupo multiplicativo das matrizes $n \times n$ não singulares sobre $\mathbf{R}$. Mostre que a aplicação $A \mapsto |A|$ é um homomorfismo de G no grupo multiplicativo dos números reais não nulos.

B.23 Seja G um grupo abeliano. Fixado $n \in \mathbf{Z}$, mostre que a aplicação $a \mapsto a^n$ é um homomorfismo de G em G.

B.24 Suponha que H e N sejam subgrupos de G, sendo N normal. Prove que $H \cap N$ é normal em H e que $H/(H \cap N)$ é isomorfo a HN/N.

Anéis

B.25 Mostre que num anel R valem as afirmações a seguir.

(i) $a \cdot 0 = 0 \cdot a = 0$, (ii) $a(-b) = (-a)b = -ab$, (iii) $(-a)(-b) = ab$.

B.26 Mostre que num anel com unidade R valem as afirmações a seguir.

(i) $(-1)a = -a$, (ii) $(-1)(-1) = 1$.

B.27 Seja R um anel e suponha que $a^2 = a$, para cada $a \in R$. Mostre que R é um anel comutativo. (Dizemos que um anel desses é um *anel booleano*.)

B.28 Seja R um anel com unidade. Fazemos de R um outro anel $\hat{R}$ definindo $a \oplus b = a + b + 1$ e $a \cdot b = ab + a + b$. (a) Verifique que $\hat{R}$ é um anel. (b) Encontre os elementos 0 e 1 de $\hat{R}$.

B.29 Seja G um grupo abeliano (aditivo) qualquer. Defina uma multiplicação em G por $ab = 0$. Mostre que isso faz de G um anel.

B.30 Demonstre o Teorema B.3. Seja I um ideal (bilateral) de um anel R. Então as classes laterais $\{a + I \mid a \in R\}$formam um anel com a adição e multiplicação de classes.

B.31 Sejam I_1 e I_2 ideais de um anel R. Prove que $I_1 + I_2$ e $I_1 \cap I_2$ também são ideais de R.

B.32 Sejam R e R' anéis. Dizemos que uma aplicação $f : R \to R'$ é um *homomorfismo* (*de anéis*) se

(i) $f(a+b) = f(a) + f(b)$ e (ii) $f(ab) = f(a)f(b)$,

para quaisquer $a, b \in R$. Prove que, se $f : R \to R'$ é um homomorfismo, então o conjunto $N = \{r \in R \mid f(r) = 0\}$ é um ideal de R. (Esse ideal N é denominado *núcleo* de f.)

Domínios e corpos

B.33 Prove que num domínio D, se $ab = ac$, $a \neq 0$, então $b = c$.

B.34 Prove que $F = \{a + b\sqrt{2} \mid a, b \text{ racionais}\}$ é um corpo.

B.35 Prove que $F = \{a + b\sqrt{2} \mid a, b \text{ inteiros}\}$ é um domínio, mas não um corpo.

B.36 Prove que um domínio D finito é um corpo.

B.37 Mostre que os únicos ideais de um corpo K são $\{0\}$ e K.

B.38 Dizemos que um número complexo $a + bi$ é um inteiro gaussiano se a, b forem inteiros. Mostre que o conjunto G dos inteiros gaussianos é um domínio. Mostre, também, que os elementos invertíveis de G são ± 1 e $\pm i$.

B.39 Sejam D um domínio e I um ideal de D. Prove que o anel quociente D/I é um domínio se, e só se, I é um ideal primo. (Dizemos que um ideal I é *primo* se $ab \in I$ implicar $a \in I$ ou $b \in I$.)

B.40 Considere o domínio $D = \{a + b\sqrt{13} \mid a, b$ inteiros$\}$ (ver Exemplo B.11.) Se $\alpha = a + b\sqrt{13}$, definimos $N(\alpha) = a^2 - 13b^2$. Prove as afirmações dadas. (i) $N(\alpha\beta) = N(\alpha)N(\beta)$; (ii) α é uma unidade se, e só se, $N(\alpha) = \pm 1$; (iii) as unidades de D são $\pm 1,\ 18 \pm 5\sqrt{13}$ e $-18 \pm 5\sqrt{13}$; (iv) os números $2,\ 3 - \sqrt{13}$ e $-3 - \sqrt{13}$ são irredutíveis.

Módulos

B.41 Seja M um R-módulo. Mostre que se A e B são submódulos de M, então $A + B$ e $A \cap B$ também são submódulos de M.

B.42 Sejam M um R-módulo e N um submódulo de M. Mostre que as classes laterais $\{u + N : u \in M\}$ formam um R-módulo com a adição de classes laterais e a multiplicação por escalar definida por $r(u + N) = ru + N$. (Esse módulo é denotado por M/N e é denominado *módulo quociente*.)

B.43 Sejam M e M' R-módulos e seja $f : M \to M'$ um R-homomorfismo. Mostre que o conjunto $N = \{u \in M : f(u) = 0\}$ é um submódulo de M. (Esse submódulo é denominado *núcleo* de f.)

B.44 Seja M um R-módulo e denote por $E(M)$ o conjunto de todos os R-homomorfismos de M em si mesmo. Defina as operações apropriadas de adição e multiplicação em $E(M)$ de tal modo que $E(M)$ se torne um anel.

Polinômios Sobre um Corpo

C.1 INTRODUÇÃO

Neste apêndice investigamos os polinômios sobre um corpo K e mostramos que eles têm muitas propriedades que são análogas às dos números inteiros. Esses resultados desempenham um papel importante na obtenção das formas canônicas de um operador linear T de um espaço vetorial V sobre K.

C.2 ANEL DE POLINÔMIOS

Seja K um corpo. Formalmente, um polinômio f sobre K é uma sequência infinita de elementos de K em que todos, exceto um número finito deles, são nulos, como segue.

$$f = (\ldots, 0, a_n, \ldots, a_1, a_0)$$

(Escrevemos a sequência de tal modo que ela se estenda para a esquerda em vez de para a direita.) O elemento a_k é denominado k-ésimo *coeficiente* de f. Se n for o maior inteiro para o qual $a_n \neq 0$, então dizemos que n é o *grau* de f e escrevemos

$$\text{grau } f = n$$

Também dizemos que a_n é o coeficiente *dominante* de f e, se $a_n = 1$, dizemos que f é um polinômio *mônico*. Por outro lado, se cada coeficiente de f for igual a 0, dizemos que f é o polinômio *nulo* ou *zero*, e escrevemos $f = 0$. Não se define o grau do polinômio nulo.

Se g for um outro polinômio sobre K, digamos,

$$g = (\ldots, 0, b_m, \ldots, b_1, b_0)$$

então a soma $f + g$ é o polinômio obtido pela soma dos coeficientes correspondentes. Isso é, se $m \leq n$, então

$$f + g = (\ldots, 0, a_n, \ldots, a_m + b_m, \ldots, a_1 + b_1, a_0 + b_0)$$

Além disso, o produto fg é o polinômio

$$fg = (\ldots, 0, a_n b_m, \ldots, a_1 b_0 + a_0 b_1, a_0 b_0)$$

ou seja, o k-ésimo coeficiente c_k de fg é dado por

$$c_k = \sum_{t=0}^{k} a_1 b_{k-1} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$$

Vale o teorema seguinte.

Teorema C.1 O conjunto P dos polinômios sobre um corpo K com as operações de adição e multiplicação forma um anel comutativo com unidade e sem divisores de zero, ou seja, é um domínio. Se f e g são polinômios não nulos de P, então grau(fg) = grau(f) grau(g).

Notação

Identificamos o escalar $a_0 \in K$ com o polinômio

$$a_0 = (\ldots, 0, a_0)$$

Também escolhemos um símbolo, digamos, t, para denotar o polinômio

$$t = (\ldots, 0, 1, 0)$$

Dizemos que t é uma *indeterminada*. Multiplicando t por si mesmo, obtemos

$$t^2 = (\ldots, 0, 1, 0, 0), \quad t^3 = (\ldots, 0, 1, 0, 0, 0), \quad \ldots$$

Assim, podemos escrever o polinômio f dado na forma tradicional

$$f = a_n t^n + \cdots + a_1 t + a_0$$

Quando escolhemos t como a indeterminada, denotamos o anel de polinômios sobre K por

$$K[t]$$

e um polinômio f costuma ser escrito como $f(t)$.

Também consideramos K como um subconjunto de $K[t]$, devido à identificação dada. Isso é possível porque as operações de adição e multiplicação de elementos de K são preservadas pela identificação, como segue.

$$(\ldots, 0, a_0) + (\ldots, 0, b_0) = (\ldots, 0, a_0 + b_0)$$
$$(\ldots, 0, a_0) \cdot (\ldots, 0, b_0) = (\ldots, 0, a_0 b_0)$$

Observamos que os elementos não nulos de K são os invertíveis do anel $K[t]$.

Também observamos que cada polinômio não nulo é um associado de um único polinômio mônico. Logo, se d e d' forem polinômios mônicos tais que d divide d' e d' divide d, então $d = d'$. (Como em todo anel, dizemos que um polinômio g divide um polinômio f se existir um polinômio h tal que $f = gh$.)

C.3 DIVISIBILIDADE

O teorema seguinte formaliza a divisão polinomial.

Teorema C.2 (Algoritmo de Divisão) Sejam f e g polinômios sobre um corpo K com $g \neq 0$. Então existem polinômios q e r tais que

$$f = qg + r$$

em que, ou $r = 0$, ou grau r < grau g.

Demonstração. Se $f = 0$ ou se grau f < grau g, então temos a representação procurada

$$f = 0g + f$$

Agora, suponha que grau $f \geq$ grau g, digamos,

$$f = a_n t^n + \cdots + a_1 t + a_0 \quad \text{e} \quad g = b_m t^m + \cdots + b_1 t + b_0$$

em que $a_n, b_m \neq 0$ e $n \geq m$. Formamos o polinômio

$$f_1 = f - \frac{a_n}{b_m} t^{n-m} g \qquad (1)$$

Então grau $f_1 <$ grau f. Por indução, existem polinômios q_1 e r tais que

$$f_1 = q_1 g + r$$

em que, ou $r = 0$, ou grau $r <$ grau g. Substituindo em (1) e resolvendo, obtemos

$$f = \left(q_1 + \frac{a_n}{b_m} t^{n-m}\right) g + r$$

que é a representação procurada.

Teorema C.3 O anel $K[t]$ de polinômios sobre um corpo K é um anel de ideais principais. Se I é um ideal de $K[t]$, então existe um único polinômio mônico d que gera I, tal que d divide cada polinômio $f \in I$.

Demonstração. Seja d um polinômio de menor grau em I. Como podemos multiplicar d por um escalar não nulo e ainda permanecer em I, podemos supor, sem perda de generalidade, que d seja um polinômio mônico. Agora, seja $f \in I$ dado. Pelo Teorema C.2, existem polinômios q e r tais que

$f = qd + r$, em que, ou $r = 0$, ou grau $r <$ grau d.

Agora, $f, d \in I$ implica $qd \in I$ e, portanto, $r = f - qd \in I$. Como d é um polinômio de grau mínimo em I, segue que $r = 0$ e $f = dq$, isto é, d divide f. Falta mostrar que d é único. Se d' for outro polinômio mônico que gera I, então d divide d' e d' divide d. Isso implica que $d = d'$, porque d e d' são mônicos. Assim, o teorema está demonstrado.

Teorema C.4 Sejam f e g polinômios não nulos de $K[t]$. Então existe um único polinômio mônico d tal que (i) d divide f e g e (ii) se d' divide f e g, então d' divide d.

DEFINIÇÃO O polinômio d do teorema precedente é denominado *máximo divisor comum* de f e g. Se $d = 1$, então dizemos que f e g são *primos entre si*.

Demonstração do Teorema C.4. O conjunto $I = \{mf + ng \mid m, n \in K[t]\}$ é um ideal. Seja d o polinômio mônico que gera I. Como $f, g \in I$, temos que d divide f e g. Suponha, agora, que d' divide f e g. Seja J o ideal gerado por d'. Então $f, g \in J$ e, portanto, $I \subset J$. Dessa forma, $d \in J$ e, portanto, d' divide d, como queríamos mostrar. Falta mostrar que d é único. Se d_1 for um outro máximo divisor comum (mônico) de f e g, então d divide d_1 e d_1 divide d. Isso implica que $d = d_1$, porque d e d_1 são mônicos. Assim, o teorema está demonstrado.

Corolário C.5 Seja d o máximo divisor comum dos polinômios f e g. Então existem polinômios m e n tais que $d = mf + ng$. Em particular, se f e g são primos entre si, então existem polinômios m e n tais que $mf + ng = 1$.

O corolário segue diretamente do fato de d gerar o ideal

$$I = \{mf + ng \mid m, n \in K[t]\}$$

C.4 FATORAÇÃO

Dizemos que um polinômio $p \in K[t]$ de grau positivo é irredutível se $p = fg$ implicar que f ou g é um escalar.

Lema C.6 Suponha que $p \in K[t]$ seja irredutível. Se p divide o produto fg de polinômios $f, g \in K[t]$, então p divide f ou p divide g. Mais geralmente, se p divide o produto de n polinômios $f_1, f_2, ..., f_n$, então p divide algum desses polinômios.

Demonstração. Suponha que p divida fg mas não f. Por ser p irredutível, os polinômios f e p devem ser primos entre si. Assim, existem polinômios $m, n \in K[t]$ tais que $mf + np = 1$. Multiplicando essa equação por g, obtemos $g = mfg + npg$. Como p divide fg e, portanto, mfg, p divide npg. Logo, p divide a soma $g = mfg + npg$.

Suponha, agora, que p divida $f_1 f_2 \cdots f_n$. Se p dividir f_1, estamos prontos. Caso contrário, pelo que acabamos de provar, p divide o produto $f_2 \cdots f_n$. Por indução em n, p divide algum dos polinômios f_2, ..., f_n. Assim, provamos o lema.

Teorema C.7 (Teorema da Fatoração Única) Seja f um polinômio não nulo de $K[t]$. Então f pode ser escrito de modo único (exceto pela ordem dos fatores) como um produto

$$f = kp_1p_2 \cdots p_n$$

em que $k \in K$ e os p_i são polinômios mônicos irredutíveis de $K[t]$.

Demonstração. Começamos demonstrando a existência de tal produto. Se f for irredutível, ou se $f \in K$, então tal produto claramente existe. Por outro lado, suponhamos que $f = gh$, onde f e g não são escalares. Então g e h têm grau menor do que o de f. Por indução, podemos supor que

$$g = k_1g_1g_2 \cdots g_r \quad \text{e} \quad h = k_2h_1h_2 \cdots h_s$$

com $k_1, k_2 \in K$ e os g_i e h_j são polinômios mônicos irredutíveis. Dessa forma,

$$f = (k_1k_2)g_1g_2 \cdots g_rk_1h_2 \cdots h_s$$

é a representação procurada.

Mostremos, agora, a unicidade (exceto pela ordem dos fatores) dessa fatoração de f. Suponhamos que

$$f = kp_1p_2 \cdots p_n = k'q_1q_2 \cdots q_m$$

com $k, k' \in K$ e os p_1, p_2, ..., p_n, q_1, q_2, ..., q_m são polinômios mônicos irredutíveis. Agora, p_1 divide $k'q_1 \cdots q_m$. Como p_1 é irredutível, o lema precedente garante que p_1 deve dividir um dos q_i.

Digamos que p_1 divida q_1. Como p_1 e q_1 são, ambos, irredutíveis e mônicos, $p_1 = q_1$. Dessa forma,

$$kp_2 \cdots p_n = k'q_2 \cdots q_m$$

Por indução, temos que $n = m$ e $p_2 = q_2$, ..., $p_n = q_m$ para algum rearranjo dos q_i. Também temos $k = k'$ e o teorema está demonstrado.

Se o corpo K for o corpo dos complexos $\mathbf{C}$, é válido o resultado seguinte, conhecido como teorema fundamental da Álgebra, cuja demonstração está fora do alcance deste texto.

Teorema C.8 (Teorema Fundamental da Álgebra) Seja $f(t)$ um polinômio não nulo sobre o corpo complexo $\mathbf{C}$. Então $f(t)$ pode ser escrito de modo único (exceto pela ordem) como um produto

$$f(t) = k(t - r_2)(t - r_2) \cdots (t - r_n)$$

em que $k, r_i \in \mathbf{C}$, isto é, como um produto de polinômios de grau 1.

No caso do corpo real $\mathbf{R}$, temos o resultado a seguir.

Teorema C.9 Seja $f(t)$ um polinômio não nulo sobre o corpo real $\mathbf{R}$. Então $f(t)$ pode ser escrito de modo único (exceto pela ordem) como um produto

$$f(t) = kp_1(t)p_2(t) \cdots p_m(t)$$

em que $k \in \mathbf{R}$ e os $p_i(t)$ são polinômios mônicos irredutíveis de grau 1 ou 2.

Miscelânea

D.1 INTRODUÇÃO

Neste apêndice discutimos vários tópicos, tais como relações de equivalência, determinantes e matrizes em bloco e a inversa generalizada de Moore-Penrose.

D.2 RELAÇÕES E RELAÇÕES DE EQUIVALÊNCIA

Uma *relação binária* R ou, simplesmente, uma *relação*, de um conjunto A para um conjunto B atribui a cada par ordenado $(a, b) \in A \times B$ exatamente uma das afirmações seguintes.

(i) "a está relacionado com b", denotado por $a\,R\,b$; (ii) "a não está relacionado com b".

Uma relação de um conjunto A para o mesmo conjunto A é denominada relação em A.

Observe que qualquer relação R de A para B define de modo único um subconjunto $\hat{R}$ de $A \times B$, como segue.

$$\hat{R} = \{(a, b) | a\,R\,b\}$$

Reciprocamente, qualquer subconjunto $\hat{R}$ de $A \times B$ define uma relação de A para B, como segue.

$$a\,R\,b \text{ se, e só se, } (a, b) \in R.$$

Em virtude da correspondência dada entre relações de A para B e subconjuntos de $A \times B$, redefinimos uma relação como segue.

DEFINIÇÃO Uma relação R de A para B é um subconjunto de $A \times B$.

Relações de equivalência

Considere um conjunto não vazio S. Uma relação R em S é denominada *relação de equivalência* em S se R for reflexiva, simétrica e transitiva, ou seja, se R satisfizer os três axiomas seguintes.

[E_1] (Reflexividade) Cada $a \in A$ está relacionado consigo mesmo, ou seja, para cada $a \in A$, temos $a\,R\,a$.
[E_2] (Simetria) Se a está relacionado com b, então b está relacionado com a, ou seja, se $a\,R\,a$, então $b\,R\,a$.
[E_3] (Transitividade) Se a está relacionado com b e b está relacionado com c, então a está relacionado com c, ou seja, se $a\,R\,b$ e $b\,R\,c$, então $a\,R\,c$.

A ideia geral motivadora do conceito de relação de equivalência é a da classificação de objetos que, de alguma forma, são "iguais". Claramente, a relação de igualdade é uma relação de equivalência. Por isso, muitas vezes usamos os símbolos ~ ou $\equiv$ para denotar uma relação de equivalência.

Exemplo D.1

(a) Na Geometria Euclidiana, a semelhança de triângulos é uma relação de equivalência. De fato, se α, β, γ são triângulos, então (i) α é semelhante a si mesmo, (ii) se α é semelhante a β, então β é semelhante a α e (iii) se α é semelhante a β e β é semelhante a γ, então α é semelhante a γ.
(b) A relação $\subseteq$ de inclusão de conjuntos não é uma relação de equivalência. A inclusão é reflexiva e transitiva, mas não é simétrica, porque $A \subseteq B$ não implica $B \subseteq A$.

Relações de equivalência e partições

Seja S um conjunto não vazio. Dizemos que uma *partição* de S é uma subdivisão de S em subconjuntos não vazios e dois a dois disjuntos, ou seja, uma coleção $P = \{A_j\}$ de subconjuntos não vazios de S tal que (i) cada $a \in S$ pertence a algum dos A_j e (ii) os conjuntos A_j são dois a dois disjuntos.

Os conjuntos de uma partição são denominados *células*. Assim, cada $a \in S$ pertence a exatamente uma das células. Também dizemos que um elemento $b \in A_j$ qualquer é um *representante* da célula A_j e que um subconjunto B de S é um *sistema de representantes* da partição se B contiver exatamente um elemento de cada uma das células da partição.

Suponha, agora, que R seja uma relação de equivalência num conjunto não vazio S. Para cada $a \in S$, a classe de equivalência de a, denotada por $[a]$, é o conjunto dos elementos de S relacionados com a, como segue.

$$[a] = \{x \mid a\,R\,x\}$$

A coleção de classes de equivalência é denominada *conjunto quociente* de S por R e denotada por S/R.

$$S/R = \{[a] \mid a \in S\}$$

A propriedade fundamental de uma relação equivalência e seu conjunto quociente é dada no teorema a seguir.

Teorema D.1 Seja R uma relação de equivalência num conjunto não vazio S. Então o conjunto quociente S/R é uma partição de S.

Exemplo D.2 Seja $\equiv$ a relação no conjunto $\mathbf{Z}$ dos inteiros definida por

$$x \equiv y \pmod 5$$

que se lê "x é congruente a y módulo 5" e que significa que a diferença $x - y$ é divisível por 5. Então $\equiv$ é uma relação de equivalência em $\mathbf{Z}$.

Existem exatamente cinco classes de equivalência no conjunto quociente $\mathbf{Z}/\equiv$, como segue.

$$\begin{aligned}
A_0 &= \{\ldots, -10, -5, 0, 5, 10, \ldots\} \\
A_1 &= \{\ldots, -9, -4, 1, 6, 11, \ldots\} \\
A_2 &= \{\ldots, -8, -3, 2, 7, 12, \ldots\} \\
A_3 &= \{\ldots, -7, -2, 3, 8, 13, \ldots\} \\
A_4 &= \{\ldots, -6, -1, 4, 9, 14, \ldots\}
\end{aligned}$$

Observe que cada inteiro x, que pode ser escrito de modo único na forma $x = 5q + r$, com $0 \le r < 5$, é um elemento da classe de equivalência A_r, em que r é o resto. Como esperamos, as classes de equivalência são disjuntas e sua união é $\mathbf{Z}$.

$$\mathbf{Z} = A_0 \cup A_1 \cup A_2 \cup A_3 \cup A_4$$

Esse conjunto quociente $\mathbf{Z}/\equiv$ é denominado conjunto dos *inteiros módulo 5* e denotado

$$\mathbf{Z}/5\mathbf{Z} \text{ ou, simplesmente, } \mathbf{Z}_5$$

Em geral, escolhemos $\{0, 1, 2, 3, 4\}$ ou $\{-2, -1, 0, 1, 2\}$ como um sistema de representantes das classes de equivalência.

Analogamente, dado qualquer inteiro positivo m, existe uma relação de equivalência $\equiv$ definida por

$$x \equiv y \pmod m$$

e o conjunto quociente $\mathbf{Z}/\equiv$ é denominado conjunto dos *inteiros módulo m*.

D.3 DETERMINANTES E MATRIZES EM BLOCO

No Capítulo 8 vimos o teorema a seguir.

Teorema 8.12 Seja M uma matriz triangular superior (inferior) em blocos com blocos diagonais $A_1, A_2, ..., A_n$. Então $\det(M) = \det(A_j)\det(A_2)\cdots\det(A_n)$.

Em vista desse resultado, se $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$, em que A é uma matriz $r \times r$ e D é $s \times s$, então $\det(M) = \det(A)$ $\det(D)$.

Teorema D.2 Considere a matriz em blocos $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, em que A é uma matriz $r \times r$ não singular e D é $s \times s$. Então $\det(M) = \det(A)\det(D - CA^{-1}B)$.

Demonstração. Decorre do fato de que $M = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$ e do teorema precedente.

D.4 FATORAÇÃO DE POSTO TOTAL

Dizemos que uma matriz B tem *posto linha total* r se B tem r linhas linearmente independentes. Dizemos que uma matriz C tem *posto coluna total* r se C tem r colunas linearmente independentes.

DEFINIÇÃO Seja A uma matriz $m \times n$ de posto r. Dizemos que uma fatoração

$$A = BC$$

de A é uma *fatoração de posto total* se B tiver posto linha total r e C tiver posto coluna total r.

Teorema D.3 Toda matriz A de posto $r > 0$ tem alguma fatoração de posto total.

Existem várias fatorações de posto total de uma matriz A. A Figura D-1 dá um algoritmo para encontrar uma dessas fatorações.

Algoritmo D.1 Dada qualquer matriz A de posto $r > 0$, este algoritmo produz uma fatoração de posto total de A.

Passo 1 Encontre a forma canônica por linhas M de A.

Passo 2 Seja B a matriz cujas colunas são as colunas de A correspondentes às colunas de M com pivôs.

Passo 3 Seja C a matriz cujas linhas são as linhas não nulas de M.

Então $A = BC$ é uma fatoração de posto total de A.

Figura D-1

Exemplo D.3 Sejam $A = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -1 & 3 \\ -1 & -1 & 2 & -3 \end{bmatrix}$ e $M = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ a forma canônica por linhas de A. Tomamos

$$B = \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Então $A = BC$ é uma fatoração de posto total de A.

D.5 INVERSA GENERALIZADA (MOORE-PENROSE)

Nesta subseção vamos supor que o corpo de escalares seja o corpo complexo **C**. Dessa forma, A^H indica a transposta conjugada de uma matriz A. [Se A for uma matriz real, então $A^H = A^T$.]

DEFINIÇÃO Seja A uma matriz $m \times n$ sobre **C**. Dizemos que uma matriz A^+ é uma pseudoinversa de A, ou inversa de Moore-Penrose de A, se $X = A^+$ satisfaz as quatro equações a seguir.

[MP$_1$] $AXA = A$,
[MP$_2$] $XAX = X$,
[MP$_3$] $(AX)^H = AX$,
[MP$_4$] $(XA)^H = XA$.

Claramente, A^+ é uma matriz $n \times m$. Além disso, se A for não singular, então $A^+ = A^{-1}$.

Lema D.4 A matriz A^+ é única (se existir).
Demonstração. Suponha que X e Y satisfaçam as quatro equações MP. Então

$$AY = (AY)^H = (AXAY)^H = (AY)^H(AX)^H = AYAX = (AYA)X = AX$$

Na primeira e quarta igualdades usamos [MP$_3$] e na segunda e última usamos [MP$_1$]. Analogamente, usando [MP$_4$] e [MP$_2$], obtemos $YA = XA$. Então,

$$Y = YAY = (YA)Y = (XA)Y = X(AY) = X(AX) = X$$

em que na primeira igualdade usamos [MP$_2$].

Lema D.5 Toda matriz A tem pseudoinversa A^+.
A Figura D-2 dá um algoritmo para encontrar a pseudoinversa de uma matriz A qualquer.

Algoritmo D.1 Dada qualquer matriz $m \times n$ A sobre **C** de posto r, este algoritmo produz a inversa de Moore-Penrose A^+ de A.

Passo 1 Troque as linhas e colunas de A de tal modo que $PAQ = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, em que A_{11} é um bloco não singular $r \times r$. [Aqui, P e Q são os produtos de matrizes elementares correspondentes às trocas de linhas e colunas.]

Passo 2 Defina $B = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$ e $C = [\mathrm{I_r}, \mathrm{A_{11}^{-1}A_{12}}]$, em que I_r é a matriz identidade $r \times r$.

Passo 3 Considere $A^+ = Q\left[C^{\mathrm{H}}(CC^{\mathrm{H}})^{-1}(B^{\mathrm{H}}B)^{-1}B^{11}\right]P$.

Figura D-2

Combinando os dois lemas precedentes, obtemos o teorema seguinte.

Teorema D.6 Toda matriz A sobre **C** tem uma única matriz A^+ de Moore-Penrose.
Quando A tiver posto linha total ou posto coluna total, temos os casos especiais a seguir.

Teorema D.7 Seja A uma matriz sobre **C**.

(a) Se A tem posto coluna total (as colunas de A são linearmente independentes), então $A^+ = (A^{\mathrm{H}}A)^{-1}$
(b) Se A tem posto linha total (as linhas de A são linearmente independentes), então $A^+ = A^H(AA^H)^{-1}$.

Teorema D.8 Sejam A uma matriz sobre **C** e $A = BC$ uma fatoração de posto total de A. Então

$$A^+ = C^+B^+ = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$$

Além disso, $AA^+ = BB^+$ e $A^+A = C^+C$.

Exemplo D.4 Considere a fatoração de posto total $A = BC$ do Exemplo D.3, ou seja,

$$A = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -1 & 3 \\ -1 & -1 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} = BC$$

Então

$$(CC^H)^{-1} = \frac{1}{5}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \; C(CC^H)^{-1} = \frac{1}{5}\begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 1 & 3 \\ 1 & -2 \end{bmatrix}, \; (B^HB)^{-1} = \frac{1}{11}\begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix}, \; B(B^HB)^{-1} = \frac{1}{11}\begin{bmatrix} 1 & 7 & 4 \\ -1 & 4 & 7 \end{bmatrix}$$

Em vista disso, a inversa de Moore-Penrose de A é dada por

$$A^+ = \frac{1}{55}\begin{bmatrix} 1 & 18 & 15 \\ 1 & 18 & 15 \\ -2 & 19 & 25 \\ 3 & -1 & -10 \end{bmatrix}$$

D.6 SOLUÇÃO DE MÍNIMOS QUADRADOS

Considere um sistema $AX = B$ de equações lineares. A solução de mínimos quadrados de $AX = B$ é o vetor de menor norma euclidiana que minimiza $\|AX - B\|_2$. Esse vetor é dado por

$$X = A^+B$$

[No caso em que A é invertível, temos $A^+ = A^{-1}$ e, portanto, $X = A^{-1}B$, que é a única solução do sistema.]

Exemplo D.5 Considere o sistema $AX = B$ a seguir.

$$\begin{aligned} x + y - z + 2t &= 1 \\ 2x + 2y - z + 3t &= 3 \\ -x - y + 2z - 3t &= 2 \end{aligned}$$

Então, usando a matriz obtida no Exemplo D.4,

$$A = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -1 & 3 \\ -1 & -1 & 2 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \quad A^+ = \frac{1}{55}\begin{bmatrix} 1 & 18 & 15 \\ 1 & 18 & 15 \\ -2 & 19 & 25 \\ 3 & -1 & -10 \end{bmatrix}$$

Em vista disso,

$$X = A^+B = (1/55)[85,\ 85,\ 105,\ -20]^T = [17/11,\ 17/11,\ 21/11,\ -4/11]^T$$

é o vetor de menor norma euclidiana que minimiza $\|AX - B\|_2$.

Lista de Símbolos

Índice

Q

R

S

T

U

V

Z